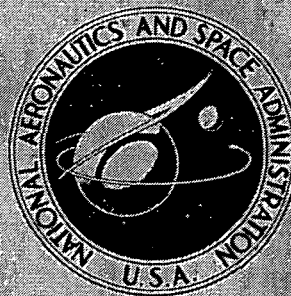


N72-27024

NASA TECHNICAL
MEMORANDUM



NASA TM X-2585

NASA TM X-2585

CASE FILE
COPY

SIMPLIFIED METHODS FOR INTERPRETING
THE EFFECT OF TRANSFER-FUNCTION ZEROS
ON THE TRANSIENT RESPONSE OF AIRCRAFT

by Reiner Onken

Ames Research Center

Moffett Field, Calif. 94035

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • JULY 1972

1. Report No. NASA TM X-2585		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle SIMPLIFIED METHODS FOR INTERPRETING THE EFFECT OF TRANSFER-FUNCTION ZEROS ON THE TRANSIENT RESPONSE OF AIRCRAFT				5. Report Date July 1972	
				6. Performing Organization Code	
7. Author(s) Reiner Onken				8. Performing Organization Report No. A-4180	
9. Performing Organization Name and Address NASA Ames Research Center Moffett Field, Calif., 94035				10. Work Unit No. 135-19-02-04	
				11. Contract or Grant No.	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D. C. 20546				13. Type of Report and Period Covered Technical Memorandum	
				14. Sponsoring Agency Code	
15. Supplementary Notes					
16. Abstract <p>Two simple methods are outlined for evaluating the effect of transfer-function zeros on the system time response. The pole effects can also be evaluated. These methods are useful for simplified analysis or creating design criteria in terms of desirable regions of pole-zero locations.</p> <p>The type of transfer function studied is limited to those of linear systems. Corresponding to ordinary longitudinal or lateral aircraft transfer functions, the denominator polynomial is of fourth order and the numerator of third order at most.</p> <p>With the longitudinal motion of the aircraft as an example, the methods are used in the evaluation of optimal regulator control with respect to a particular performance index structure.</p>					
17. Key Words (Suggested by Author(s)) Control Dynamic characteristics Roots of equations				18. Distribution Statement Unclassified - Unlimited	
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 44	
				22. Price* \$3.00	

TABLE OF CONTENTS

	Page
SYMBOLS	v
SUMMARY	1
INTRODUCTION	1
RESIDUE INTERPRETATION METHOD	2
Transfer Function and Transient Response Residues	2
Phase Loci of the Residues	4
Magnitude Loci of the Residues	7
Interpretation of Pole Modifications	11
APPROXIMATE MODAL EXPANSION METHOD	12
Modal-Type Transfer Function Expansion (Two Zeros)	12
Zero Loci Due to the β Values	13
Three Zeros	17
EVALUATION OF OPTIMAL LONGITUDINAL CONTROL TIME RESPONSE WITH USE OF POLE-ZERO INTERPRETATION METHODS	19
Problem Formulation	19
Performance Criterion	20
Discussion of Resulting Optimal Transfer Functions	21
CONCLUDING REMARKS	32
APPENDIX A - EQUATION FOR THE MAGNITUDE LOCI OF C_{iz} (TWO ZEROS)	33
APPENDIX B - ENCIRCLING OF ZERO REGIONS FOR GIVEN VALUES OF $ C_{iz} $. . .	34
Region I	35
Region II	35
Region III	37
REFERENCES	38

SYMBOLS

a_i	coefficients of the transfer-function denominator
b_i	coefficients of the transfer-function numerator
C_i	residues
C_{ip}	pole-dependent part of C_i
C_{iz}	zero effected part of C_i
c_v	portions of factorized C_{iz} due to the zero z_v
$D(s)$	transfer-function denominator (with $K = 1$)
$F(s)$	transfer function
$F = [f_{ij}]$	system matrix
$G = [g_{ij}]$	control matrix
$H = [h_{ij}]$	transfer-function matrix for response to initial conditions
I	unity matrix
J	performance criterion
K	steady-state gain
$L = [l_{ij}]$	regulator gain matrix
$N(s)$	transfer-function numerator
p_i	transfer-function poles
q	weighting of the state in performance index
r	weighting of the controls in performance index
r	radius
s	Laplace variable
t	time variable
u	error in forward velocity ($= \hat{u} \cdot V_c$)
u_c	change in control variable

V_c	command forward speed (equilibrium speed)
w	error in vertical velocity
x	state variable
z_i	transfer-function zeros
$\beta_{I,II,III}$	transfer-function expansion coefficients
ϵ_D	damping angle
δ_e	change in elevator control deflection
δ_F	change in direct lift control deflection
δ_T	change in thrust control deflection
φ_1, φ_2	angles, defined in figures 2 and 4
φ_{Nz}	phase-angle contribution of the C_{iz} denominator
ϕ_i	phase angle of C_{iz}
γ	error in glide-path angle
θ	error in longitudinal attitude (pitch angle)
σ_i	real part of p_i (damping)
σ_I	real part of poles $p_{1,2}$ (damping of eigenmode I)
σ_z	real part of pair of zeros
σ_{zm}	coordinate of the centers of zero loci circles on the real axis
ω_e	eigenmode frequency
ω_{oI}	natural frequency of eigenmode I
ω_{oz}	natural frequency of pair of zeros
ω_z	"eigenfrequency" of conjugate complex pair of zeros
$\bar{\omega}_z$	"eigenfrequency" of pair of real zeros

SIMPLIFIED METHODS FOR INTERPRETING THE EFFECT
OF TRANSFER-FUNCTION ZEROS ON THE
TRANSIENT RESPONSE OF AIRCRAFT

Reiner Onken*

Ames Research Center

SUMMARY

Two simple methods are outlined for evaluating the effect of transfer-function zeros on the system time response. The pole effects can also be evaluated. These methods are useful for simplified analysis or creating design criteria in terms of desirable regions of pole-zero locations.

The type of transfer function studied is limited to those of linear systems. Corresponding to ordinary longitudinal or lateral aircraft transfer functions, the denominator polynomial is of fourth order and the numerator of third order at most.

With the longitudinal motion of the aircraft as an example, the methods are used in the evaluation of optimal regulator control with respect to a particular performance index structure.

INTRODUCTION

The dynamic behavior of a flight vehicle usually can be described by a system of linearized differential equations with constant coefficients or by the corresponding set of transfer functions. The use of feedback control in multivariable systems makes it possible to modify not only the denominator of the transfer functions, but also the numerator. This property is desirable for it leads to the additional possibility of improving the vehicle's dynamic behavior. As an extreme case, nonminimum phase effects (numerator roots with positive real component), which occur rather often in aircraft transfer functions, could be ruled out or at least be moderated by appropriate feedback control.

Therefore, some work was directed toward the development of methods for the investigation of the numerator roots based on illustrative frequency domain techniques (refs. 1-4). These techniques provide an estimate of changes in the numerator and denominator roots of the transfer functions as a consequence of feedback loop gains. To establish design criteria in terms of appropriate root locations, the effect of the zero locations on the dynamic behavior, that is, time response, should be known. The time response appears

*National Research Council Associate.

to be most suitable for evaluations of system behavior, since the necessity of an analytical formulation of the performance index can be avoided. System time response may be obtained by resolving the transfer functions into partial fraction series and taking the inverse term by term (residue calculation). This procedure is well known, but it does not offer a means of showing transparently what changes in the time response can be expected from certain pole-zero modifications. Especially for multivariable systems, a more effective procedure is required. Various pole-zero configurations have to be evaluated to determine the effect of modifications on the corresponding time histories.

This report discusses simple methods that show the effect of zeros and poles on the transient response so that root locations can be assigned for certain desirable transient response behavior. Two basic methods will be described: The first is based on the assignment of phase and magnitude of the residues to the root locations; the second, which is easily applicable to aircraft control systems, is based on a specific expansion of the transfer functions, where each expansion term itself has a known transient response.

The application of both methods is restricted to a controlled element transfer function with a fourth-order denominator and at most third-order numerator. This typifies either the longitudinal or lateral directional equations of motion of an aircraft.

Implementation of the methods is discussed for the evaluation of an aircraft optimal control system. The optimal longitudinal motion regulator with respect to a particular quadratic performance index is evaluated for different relative weighting between the states and the controls.

RESIDUE INTERPRETATION METHOD

Transfer Function and Transient Response Residues

The transfer functions investigated here are of the fourth-order denominator and at most third-order numerator type. This is the structure of the main open-loop aircraft longitudinal motion transfer functions as well as the corresponding closed-loop transfer functions with constant feedback gains. If u_c is generally taken as the input and x as the output, the transfer function becomes

$$F(s) = \frac{x(s)}{u_c(s)} = K \frac{(-p_1)(-p_2)(-p_3)(-p_4)}{(s-p_1)(s-p_2)(s-p_3)(s-p_4)} \frac{(s-z_1)(s-z_2)(s-z_3)}{(-z_1)(-z_2)(-z_3)} \quad (1)$$

Disregarding the case of multiple roots, in the frequency domain the response to a step input can be written as a partial fraction expansion:

$$x(s) = \sum_{i=0}^4 \frac{C_i}{s - p_i} \quad \text{with } p_0 = 0 \quad (2)$$

The inverse is

$$x(t) = \sum_{i=0}^4 C_i e^{p_i t} \quad (3)$$

The partial fraction expansion coefficients C_i , weighting the different modes involved, are

$$C_i = \lim_{s \rightarrow p_i} [(s - p_i)x(s)] \quad (4)$$

Thus

$$C_i = K \frac{\prod_{v=1}^3 [(p_i - z_v)/(-z_v)]}{\prod_{\substack{u=1 \\ u \neq i}}^4 [(p_i - p_u)/(-p_u)]} \quad (5)$$

Normalizing the transient response with respect to the steady state ($K = 1$) gives $C_0 = 1$. The remaining residues become

$$C_i = - \prod_{\substack{u=1 \\ u \neq i}}^4 \left(\frac{-p_u}{p_i - p_u} \right) \cdot \prod_{v=1}^3 \left(\frac{p_i - z_v}{-z_v} \right) = -C_{ip} \cdot C_{iz} \quad (6)$$

where

$$C_{ip} = \prod_{\substack{u=1 \\ u \neq i}}^4 \frac{-p_u}{p_i - p_u} \quad \text{and} \quad C_{iz} = \prod_{v=1}^3 \frac{p_i - z_v}{-z_v}$$

The entire residue expression has been separated into two main factors C_{ip} and C_{iz} , similar to each other, where one part (C_{ip}) consists only of pole terms, while the other part (C_{iz}) contains all zero-affected terms. Thus, C_{iz} covers the whole effect of the zeros on the residue.

The zero-affected portion C_{iz} consists of products of fractions whose factors are complex numbers in terms of the poles and zeros. They could be drawn in the root plane as vectors from the different zeros to the pole p_i and the origin. For two zeros, four vectors would have to be investigated for each residue. This is not convenient, and a different representation of the zero effect on the residues is desirable. Furthermore, the representation should illuminate zero regions of similar effect. Such a representation is developed in the following sections.

The pole-dependent portion C_{ip} has a reciprocal structure to that of C_{iz} . Thus, the influence of the poles on the residues can be considered in a way similar to that of the zeros.

Once the effect of the zeros and poles on the residues is known, the time vector presentation is a helpful tool for putting together the various portions of equation (3) to determine the overall transient response (refs. 4 and 5). There are two oscillatory eigenmodes for aircraft longitudinal control: the low-frequency phugoid mode and the fast short-period mode. Consequently,

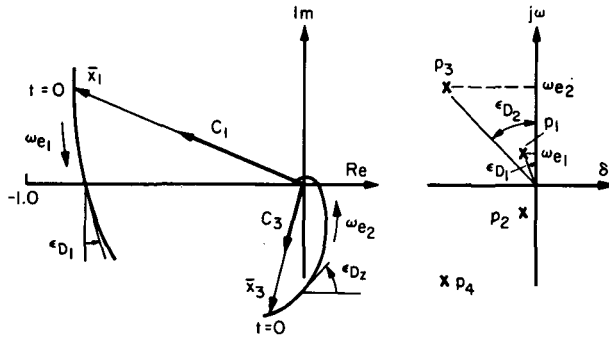


Figure 1.- Residues and time vectors for given pole combination.

there are two time vectors \bar{x}_1 and \bar{x}_3 at the time $t = 0$ (see fig. 1) in phase with C_1 and C_3 and with a magnitude of twice C_1 and C_3 .¹ They rotate with time corresponding to their eigenmode frequency, while decreasing in magnitude as a result of their eigenmode damping. The transient response for any time t , due to the residues C_1 through C_4 in equation (3), can be read from figure 1, as the sum of the real portions of the time vectors. Thus, the purpose here is to determine the change of the residues (time vectors) in phase and magnitude introduced by the zero-dependent numerators of the transfer functions.

Phase Loci of the Residues

As already indicated, the residues must be considered in general as conjugate complex quantities, and they must be described by their phase and magnitude for later use with the time vector representation. Therefore, the link between the zero location and the residues is required before the overall correspondence between the zero locations and certain desirable transient response behavior can be established. Thus, the phase contribution to the residues due to the zero locations is presented here in terms of phase loci in the root plane for the following cases:

1. Pair of conjugate complex zeros
2. Pair of real zeros
3. Three zeros

Pair of conjugate complex zeros—The portion of the residue C_1 that is influenced by the zeros is known from equation (6) as

$$C_{iz} = \frac{|p_1 - z_1| |p_1 - z_2|}{|-z_1| |-z_2|} e^{j\phi_1} \quad (7)$$

¹The residues C_2 and C_4 are implicitly included in \bar{x}_1 and \bar{x}_3 . They do not appear explicitly, because they are given as the conjugate complex numbers C_1 and C_3 .

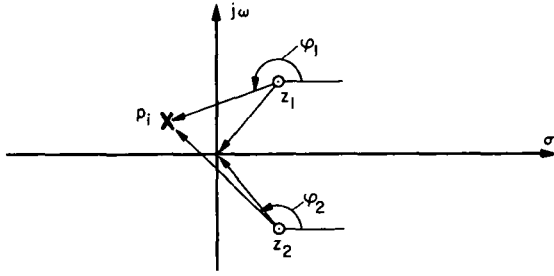


Figure 2.- Definition of φ_1 and φ_2 .

The phase angle ϕ_i represents the change in phase due to the zeros. From the four vectors involved, as shown in figure 2, the phase angle ϕ_i becomes

$$\phi_i = \varphi_1 + \varphi_2 - \varphi_{Nz} = \varphi_1 + \varphi_2 \quad (8)$$

with $\varphi_{Nz} = 0$ for pairs of zeros. To determine the loci of zeros that have equal phase changes, consider

$$\tan \phi_i = \text{const} = \frac{\tan \varphi_1 + \tan \varphi_2}{1 - \tan \varphi_1 \cdot \tan \varphi_2} \quad (9)$$

If the definition of φ_1 and φ_2 is substituted from figure 2 into equation (9), it can be shown that those loci are circles of the form

$$(\sigma_z - \sigma_{zm})^2 + \omega_z^2 = r^2 \quad (10)$$

where

$$\sigma_{zm} = \sigma_i - \frac{\omega_i}{\tan \phi_i}$$

$$r^2 = \omega_i^2 \left(1 + \frac{1}{\tan^2 \phi_i} \right)$$

$$z_{1,2} = \sigma_z \pm j\omega_z$$

$$p_i = \sigma_i + j\omega_i$$

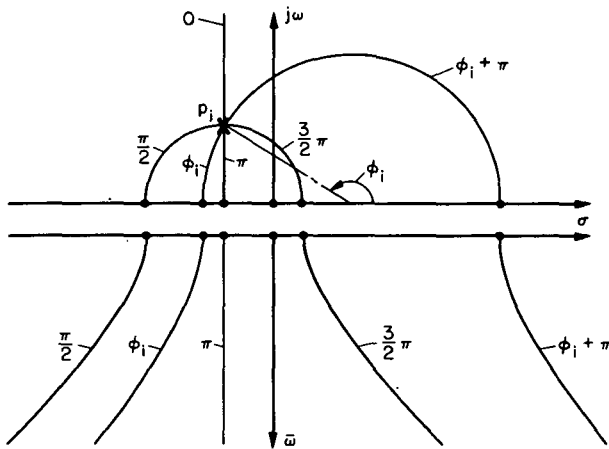


Figure 3.- Phase loci of C_{iz} .

These loci are illustrated in the upper half plane in figure 3, where it becomes evident that all circles intersect the pole p_i and have their center on the real axis. They degenerate to a vertical straight line through p_i if ϕ_i is either $0, 2\pi, \dots$, or $\pi, 3\pi, \dots$. If the portions of the circles on the left side of p_i correspond to the phase angle ϕ_i , the right-hand portions correspond to $\phi_i + \pi$. For each circle, the angle ϕ_i can easily be

determined as the angle between the real axis and the line connecting the center of the circle with p_i .

Thus, each pair of conjugate complex zeros lying on one of the circles causes the same phase shift ϕ_i on the residue C_i as any other pair of zeros on the same circle. The intersection points of the circles with the real axis correspond to double zeros. This yields the link between the case of conjugate complex zeros and that one of a pair of real zeros.

Pair of real zeros—The phase angle ϕ_i for the case of real zeros is also given by equation (8) (see fig. 4). The loci of pairs of real zeros satisfying the condition of equal phase changes are

$$(\sigma_z - \sigma_{zm})^2 - \bar{\omega}_z^2 = r^2 \quad (11)$$

where

$$\sigma_{zm} = \sigma_i - \frac{\omega_i}{\tan \phi_i}$$

$$r^2 = \omega_i^2 \left(1 + \frac{1}{\tan^2 \phi_i} \right)$$

$$z_{1,2} = \sigma_z \pm \bar{\omega}_z$$

$$p_i = \sigma_i + j\omega_i$$

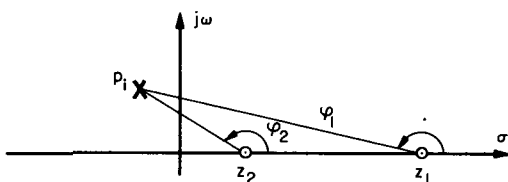


Figure 4.— Definition of ϕ_1 and ϕ_2 for pair of real zeros.

Drawing it in a $\sigma_z, \bar{\omega}_z$ plane (the lower half plane in fig. 3), we obtain a set of rectangular hyperbolas with their asymptotes inclined by 45° to the real axis. For corresponding values of ϕ_i , the hyperbolas intersect the real axis at the same points as the circles for the case of conjugate complex pairs of zeros. Thus, considering both the base of conjugate complex zeros and that of a pair of real zeros, a certain range in the

residue phase angle will correspond to an area in the root plane, bounded by the appropriate loci curves.

Three zeros—If the number of zeros is extended from two to three, the equation for the residue portion affected by the zeros C_{iz} is slightly changed to

$$C_{iz} = e^{\phi_i} \cdot \prod_{v=1}^3 \frac{|p_i - z_v|}{|-z_v|} \quad (12)$$

Considering the phase loci, there is the same family of curves in the root plane as already derived, with an additional phase shift $\Delta\phi_i$ due to the third zero; that is,

$$\phi_i = \phi_i(\text{two zeros}) + \Delta\phi_i \quad (13)$$

where $\Delta\phi_i$ can be easily derived for the third zero moving along the real axis, as shown in figure 5. There is a discontinuity in $\Delta\phi_i$ of π , when z_3 passes the imaginary axis.

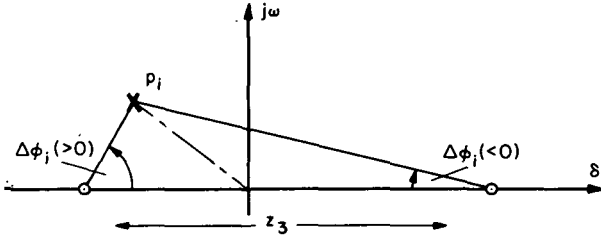


Figure 5.- Phase shift $\Delta\phi_i$ due to third zero z_3 .

Magnitude Loci of the Residues

As outlined previously, a mapping of the magnitude, as well as the phase, of C_i in the root plane by means of locus curves is required. Following the same line of reasoning as for the phase effect of the zeros on C_{iz} , one obtains locus curves of higher order, which are too complicated for simple geometrical interpretation (see appendix A). Therefore, the expression for $|C_{iz}|$ will be split so that it can be discussed partially by simple geometrical means. Again, the cases of conjugate complex pairs of zeros, two real zeros, and three zeros are discussed.

Pair of conjugate complex zeros—Following equation (7), the magnitude of C_{iz} is split into

$$|C_{iz}| = c_1 \cdot c_2 \quad (14)$$

where

$$c_1 = \frac{|p_i - z_1|}{|-z_1|}, \quad c_2 = \frac{|p_i - z_2|}{|-z_2|}$$

Therefore, the two portions c_1 and c_2 are discussed separately and then their product $|C_{iz}|$. The loci curves in the root plane for both c_1 and c_2 again are circles (fig. 6) given, respectively, by

$$(\sigma_z - \sigma_{zm1})^2 + (\omega_z - \omega_{zm1})^2 = r_{z1}^2 \quad (15)$$

and

$$(\sigma_z - \sigma_{zm2})^2 + (\omega_z - \omega_{zm2})^2 = r_{z2}^2 \quad (16)$$

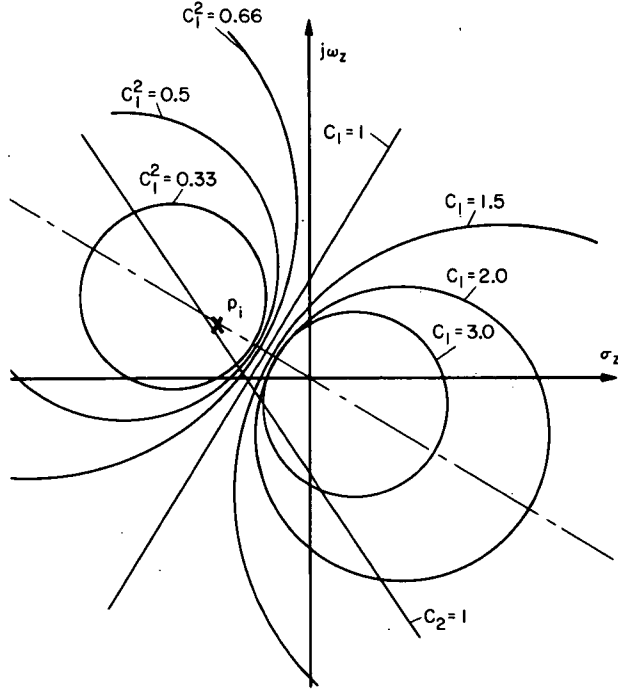


Figure 6.- Magnitude loci of c_1 .

where

$$\sigma_{zm1} = \frac{\sigma_i}{1 - c_1^2}, \quad \sigma_{zm2} = \frac{\sigma_i}{1 - c_2^2}$$

$$\omega_{zm1} = \frac{\omega_i}{1 - c_1^2}, \quad \omega_{zm2} = -\frac{\omega_i}{1 - c_2^2}$$

$$r_{z1}^2 = (\sigma_i^2 + \omega_i^2) \frac{c_1^2}{(1 - c_1^2)^2}, \quad r_{z2}^2 = (\sigma_i^2 + \omega_i^2) \frac{c_2^2}{(1 - c_2^2)^2}$$

Both families of circles are very similar: the circles for c_2 represent the symmetrical image about the real axis of those with the same value for c_1 . Thus, only circles for c_1 are shown in figure 6; the straight line representing $c_2 = 1$ is added to show the symmetry.

The geometrical structure of the circles becomes obvious by considering the following properties (stated for the case of c_1 only):

1. The center points of the circles given by σ_{zm1} and ω_{zm1} lie on the straight line passing through p_i and the origin.

2. The radius of the locus circles for $c_1 = a$ and $c_1 = 1/a$ is the same. The corresponding center points lie symmetrical to the point midway between p_i and the origin. The midpoint is on the locus circle for $c_1 = 1$, which has degenerated to a straight line and intersects the real axis at

$$\sigma_z = |p_i|^2 / 2\sigma_i \quad (17)$$

The line for $c_1 = 1$ is thus the perpendicular bisector of the line between p_i and the origin.

3. The circles degenerate to a point at p_i or at the origin for values of c_1 going to either zero or infinity, respectively.

4. The radii of the circles are given by the product between the distance of the center from the origin and the value of the corresponding magnitude.

Considering the intersections of the c_1 and c_2 circles, each such intersection represents a value of $|C_{iz}|$, given by the product of c_1 and c_2 at that point. Vice versa, for a given point in the root plane, c_1 and c_2 can be determined and then $|C_{iz}|$ itself.

Although the higher order loci curves of the magnitude of C_{iz} are given by the intersections of two families of circles, it still is rather tedious to derive those curves for each required case without great computational effort.

A simpler interpretation can be obtained by encircling an area in the root plane for zeros that correspond to the required ranges of $|C_{iz}|$. That means the exact locus curve for the required value of $|C_{iz}|$ will be approximated by a surrounding region.

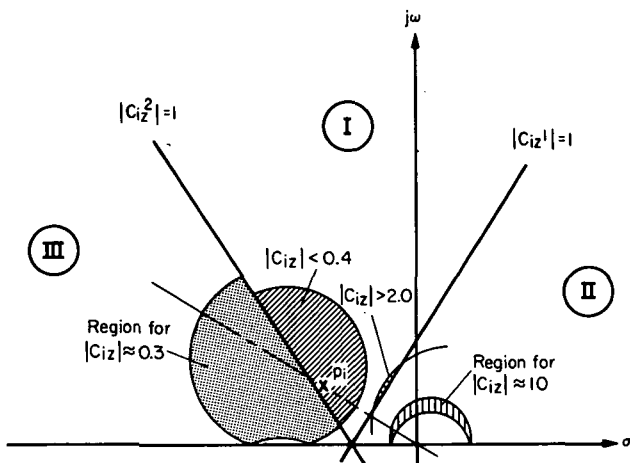


Figure 7.- Examples of approximate zero locations for certain values of $|C_{iz}|$.

Figure 7 shows the upper half of the root plane (the lower half would show symmetrically the same), where for each of the three regions (I, II, III) separated by the locus curves for $c_1 = 1$ and $c_2 = 1$, an example of encircling is shown. Those examples are based on the following statements. (More details are given in appendix B.)

Region I: (a) If $|C_{iz}|$ is distinctively greater than 1, the circle of $c_2 = |C_{iz}|$ surrounds the area within which all conjugate complex zeros corresponding to that $|C_{iz}|$ must lie (example, $|C_{iz}| > 2.0$). (b) If $|C_{iz}|$ is distinctively smaller than 1, the circle of $c_1 = |C_{iz}|$

surrounds the area within which all conjugate complex zeros corresponding to that $|C_{iz}|$ must lie (example, $|C_{iz}| < 0.4$).

Region II: All conjugate complex zeros corresponding to a certain value of $|C_{iz}| > 1$ must lie within the union of the circles for $c_1^2 = |C_{iz}|$ and $c_2^2 = |C_{iz}|$ but outside the intersection of those circles (example, $|C_{iz}| \approx 10$).

Region III: All conjugate complex zeros corresponding to a certain value of $|C_{iz}|$ where

$$1 - (\sigma_i/|p_i|)^2 < |C_{iz}|^2 < 1$$

must lie within the union of the circles for $c_1^2 = |C_{iz}|$ and $c_2^2 = |C_{iz}|$ but outside the intersection of those circles (example, $|C_{iz}| \approx 0.3$).

As the straight-line loci for $c_1 = 1$ and $c_2 = 1$ separating the different regions change with p_i , the circles encircling the zero regions are cut differently by those lines. However, the radii of the circles are proportional to $|p_i|$ and do not change if p_i remains constant in magnitude.

For the locus of $|C_{iz}| = 1$, some other simplifying specifications can be made. In the relations given in appendix A, it becomes evident that the locus for $|C_{iz}| = 1$ intersects the real axis at the same point that the loci for $c_1 = 1$ and $c_2 = 1$ intersect. The locus goes to infinity in its imaginary coordinate for

$$\sigma_z = \frac{\sigma_i^2 - \omega_i^2}{2\sigma_i} \quad (18)$$

and at $\sigma_z = \sigma_i$, ω_z is defined by

$$\omega_z^2 + \sigma_i^2 = \frac{1}{2} |p_i| \quad (19)$$

Notice that equation (19) can hold only if the locus intersects the real axis at the left of the pole p_i . Otherwise, the relation for ω_z at $\sigma_z = (1/2)\sigma_i$ may be used, where

$$\omega_z^2 + \frac{1}{4} \sigma_i^2 = \frac{1}{2} \omega_i^2 \quad (20)$$

Real zeros—Following the procedure worked out for the case of a pair of conjugate complex zeros (see fig. 6), it becomes evident that the intersection points of the circles for $c_1 = \text{const}$ with the real axis represent locations for single zeros with $|C_{iz}| = c_1$. The same is true for those circles of $c_2 = \text{const}$, which leads to the same zeros and values of $|C_{iz}|$.

For a single zero, $|C_{iz}|$ can be derived from the same set of curves used for the conjugate complex zeros, and no further set of curves for real zeros is needed. Real zeros thus are simply treated using a plot like figure 6 separately for each.

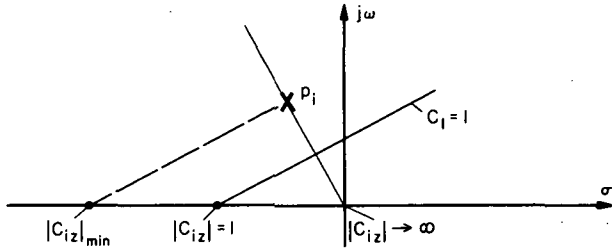


Figure 8.- Illustration of change in $|C_{iz}|$ with variation of one real zero.

How $|C_{iz}|$ changes as a consequence of shifting a simple zero along the real axis is further illustrated by figure 8. As in figure 6, the zero location for $|C_{iz}| = 1$ must lie at the point where the line for $c_1 = 1$ intersects the real axis (eq. (17)). From there $|C_{iz}|$ increases toward infinity as the zero moves toward the origin. As the zero moves beyond the origin, $|C_{iz}|$ decreases monotonically to 1.

As the zero moves in the opposite direction, the magnitude of C_{iz} decreases to a minimum value and approaches 1 again beyond the minimum point. The least achievable value of $|C_{iz}|$ is dependent only on p_i ,

$$|C_{iz}|_{\min} = 1 - \left(\frac{\sigma_i}{|p_i|} \right)^2 \quad (21)$$

and the corresponding zero location is

$$z_v = \frac{|p_i|^2}{\sigma_i}$$

which is twice the value for the intersection point of the line of $c_1 = 1$ (fig. 8).

Interpretation of Pole Modifications

The preceding discussion is concerned with C_{iz} , the zero-dependent part of the residues (eq. (6)), and the effect of zero changes on that part. The same concept can be used for the interpretation of C_{ip} , the other part of the residues (eq. (6)), which changes only with modifications of the poles. Since the transfer function denominator is restricted to fourth order, the expression

$$C_{ip} = \prod_{\substack{u=1 \\ \neq i}}^4 \frac{-pu}{p_i - pu}$$

can formally be discussed as that which would have been obtained as the reciprocal of C_{iz} with three zeros, given the values of the poles p_u . In the case of complex p_i , the remaining three poles are used, in the same manner as before. However, since the "third zero" is the pole conjugate to p_i , its phase and magnitude are constant. The direct analog to the case of C_{iz} with a pair of zeros, from which the loci plots are formed, is C_{ip}^* , the part of C_{ip} that does not contain the conjugate pole of p_i . Thus, the loci plots derived for C_{iz} can be used for C_{ip}^* as well, by relabeling the loci curves in a simple manner:

$$\phi_i[C_{iz}] \triangleq -\phi_i[C_{ip}^*] \quad (22)$$

and

$$|C_{iz}| \triangleq \frac{1}{|C_{ip}^*|} \quad (23)$$

APPROXIMATE MODAL EXPANSION METHOD

Modal-Type Transfer Function Expansion (Two Zeros)

The residue interpretation method discussed above is very appropriate for the interpretation of the effect of zeros on the time response, if the poles of the system under consideration are conjugate complex. In this case, with a fourth-order system, there are only two residues or time vectors that must be considered, each representing one of the eigenmodes. If instead of the conjugate complex poles there are real ones, the number of independently moving time vectors increases (up to at most four). Pursuing the motion of these time vectors to determine their superposition at any time t is no longer simple enough. Even in the case of three independent time vectors the method loses much of its power. Therefore, in cases where some of the poles are real, a different method is proposed based on a modal-type transfer function expansion using typical pole locations obtained from longitudinal aircraft equations of motion. According to the transfer function given in equation (1) and its later normalization, the expansion is written for the case of a pair of zeros ($b_3 = 0$):

$$\begin{aligned} F(s) &= \frac{N(s)}{D(s)} = \frac{1+b_1s+b_2s^2}{1+a_1s+a_2s^2+a_3s^3+a_4s^4} \\ &= \frac{\beta_I}{1+\alpha_{1I}s+\alpha_{2I}s^2} + \frac{\beta_{II}}{1+\alpha_{1II}s+\alpha_{2II}s^2} + \frac{\beta_{III}}{(1+\alpha_{1I}s+\alpha_{2I}s^2)(1+\alpha_{1II}s+\alpha_{2II}s^2)} \\ &= \frac{\beta_I}{D_I(s)} + \frac{\beta_{II}}{D_{II}(s)} + \frac{\beta_{III}}{D(s)} \end{aligned} \quad (24)$$

The denominator of the third expansion term is that of the original transfer function, and the denominators of the two other terms are the characteristic polynomials of the eigenmodes of the aircraft called the phugoid (β_I) and the short-period motion (β_{II}). These denominators will not be changed if one pair of the poles or even all poles become real. The eigenmodes are considered to be known in advance as their response characteristics. Thus the gains β_I and β_{II} represent a kind of modal expansion, and the β_{III} term should be incorporated in that expansion. With regard to the aircraft longitudinal motion, this can be done by an approximation (ref. 5). It will be assumed that β_{III}/D can be expressed using only the phugoid portion of the denominator. This is the case if (1) the values of β_I and β_{III} are not very high and (2) the magnitude ratio between the resulting time vector of the supposedly aperiodic phugoid and that of the short period, both at time $t = 0$, is greater than about 10. Usually the magnitude ratio $|p_3 - p_2|/|p_1 - p_4|$ is about 1, so the condition for the time vector ratio is met, if

$$|p_3| > 3\sqrt{p_1 \cdot p_2} \quad (25)$$

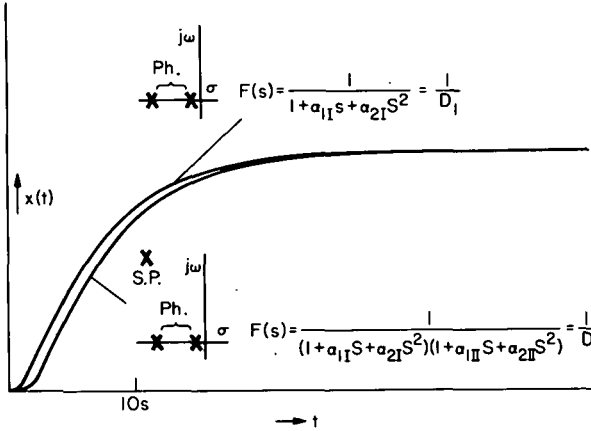


Figure 9.- Comparison between actual $F(s)$ and second-order approximation.

Figure 9 gives some evidence of the degree of coincidence that may be expected between the first and third expansion term time histories.

With these assumptions, the expansion of the transfer function in equation (24) can be modified into

$$F(s) \approx \frac{\beta_I + \beta_{III}}{1 + \alpha_{1I}s + \alpha_{2I}s^2} + \frac{\beta_{II}}{1 + \alpha_{1II}s + \alpha_{2II}s^2} \quad (26)$$

By this type of modal expansion, the superposition of only two modally characteristic time histories has to be made. The gain $(\beta_I + \beta_{III})$ on the one hand and β_{II} on the other hand give the degree of dominance of one of the two portions.

Zero Loci Due to the β Values

To assign zero locations to certain desired types of superposition in phugoid and short-period response given by β_I , β_{II} , and β_{III} , zero loci for constant β values can be plotted in the root plane. Equation (24) yields

$$N(s) = 1 + b_1 s + b_2 s^2 = \beta_I (1 + \alpha_{1II} s + \alpha_{2II} s^2) + \beta_{II} (1 + \alpha_{1I} s + \alpha_{2I} s^2) + \beta_{III} \quad (27)$$

so that β_I , β_{II} , and β_{III} are given as quantities dependent on b_1 and b_2 , that is, on the zero locations, while the poles are considered as fixed:

$$\left. \begin{aligned} \beta_I &= \frac{\omega_{oII}^2}{\omega_{oz}^2} \cdot \frac{\sigma_z - \sigma_I}{\sigma_{II} - \sigma_I} \\ \beta_{II} &= - \frac{\omega_{oI}^2}{\omega_{oz}^2} \cdot \frac{\sigma_z - \sigma_{II}}{\sigma_{II} - \sigma_I} \\ \beta_{III} &= 1 - (\beta_I + \beta_{II}) \end{aligned} \right\} \quad (28)$$

with

$$\begin{aligned} \sigma_z &= - \frac{b_1}{2b_2} \\ \omega_{oz}^2 &= \frac{1}{b_2} \\ \sigma_I &= - \frac{\alpha_{1I}}{2\alpha_{2I}} \\ \sigma_{II} &= - \frac{\alpha_{1II}}{2\alpha_{2II}} \\ \omega_{oI}^2 &= \frac{1}{\alpha_{2I}} \\ \omega_{oII}^2 &= \frac{1}{\alpha_{2II}} \end{aligned}$$

These equations represent the loci curves. When written in terms of σ_z and ω_z for the case of conjugate complex pairs of zeros, they again are circles (fig. 10):

$$\left. \begin{aligned} \left[\sigma_z - \frac{\omega_{oII}^2}{2\beta_I(\sigma_{II} - \sigma_I)} \right]^2 + \omega_z^2 &= \frac{\omega_{oII}^2}{2\beta_I(\sigma_{II} - \sigma_I)} \left[\frac{\omega_{oII}^2}{2\beta_I(\sigma_{II} - \sigma_I)} - 2\sigma_I \right] \\ \left[\sigma_z + \frac{\omega_{oI}^2}{2\beta_{II}(\sigma_{II} - \sigma_I)} \right]^2 + \omega_z^2 &= \frac{\omega_{oI}^2}{2\beta_{II}(\sigma_{II} - \sigma_I)} \left[\frac{\omega_{oI}^2}{2\beta_{II}(\sigma_{II} - \sigma_I)} + 2\sigma_{II} \right] \end{aligned} \right\} \quad (29)$$

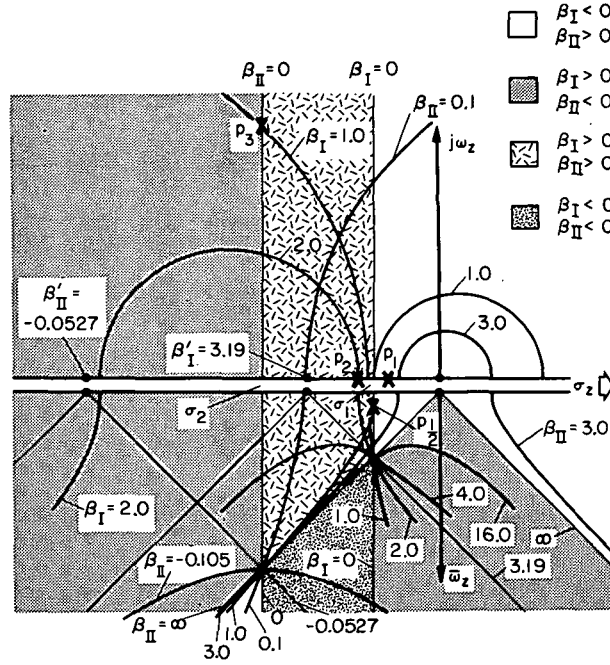


Figure 10.- β_I and β_{II} loci.

with $\omega_{OZ}^2 = \sigma_Z^2 + \bar{\omega}_Z^2$. In case of pairs of real zeros with root plane axes for σ_Z and $\bar{\omega}_Z$, and

$$\omega_{OZ}^2 = \sigma_Z^2 - \bar{\omega}_Z^2 \quad (30)$$

they become rectangular hyperbolas of the same algebraic structure:

$$\left\{ \begin{aligned} \left[\sigma_Z - \frac{\omega_{OII}^2}{2\beta_I(\sigma_{II} - \sigma_I)} \right]^2 - \bar{\omega}_Z^2 &= \frac{\omega_{OII}^2}{2\beta_I(\sigma_{II} - \sigma_I)} \left[\frac{\omega_{OII}^2}{2\beta_I(\sigma_{II} - \sigma_I)} - 2\sigma_I \right] \\ \left[\sigma_Z + \frac{\omega_{OI}^2}{2\beta_{II}(\sigma_{II} - \sigma_I)} \right]^2 - \bar{\omega}_Z^2 &= \frac{\omega_{OI}^2}{2\beta_{II}(\sigma_{II} - \sigma_I)} \left[\frac{\omega_{OI}^2}{2\beta_{II}(\sigma_{II} - \sigma_I)} + 2\sigma_{II} \right] \end{aligned} \right\} \quad (31)$$

With respect to β_I (or β_{II}) all hyperbolas intersect in one point, geometrically given by the intersection of the 45° line through the origin and the line $\sigma_Z = \sigma_1$ (or σ_2). Figure 10 shows a plot of some loci curves for both conjugate complex (upper) and real (lower) pairs of zeros. The phugoid poles are fixed at $p_1 = -0.186$ (1/sec), $p_2 = -0.3$ (1/sec), and the short-period poles at $p_{3,4} = -0.65 \pm j 0.92$ (1/sec). For simplicity, only loci for positive values of β_I and β_{II} are shown. The β_{III} loci are also not plotted, because the value of β_{III} can easily be determined from β_I and β_{II} (eq. (28)). If the pair of zeros is located on either the phugoid or the short-period poles, β_{III} becomes zero. According to the approximation made in the

last section, the loci circles of constant β_{III} are almost coincident with those of β_I , but with different values assigned to them corresponding to equation (28). In figure 10, the shading indicates regions of zero locations for which certain combinations in the signs of β_I and β_{II} are valid. It appears, for example, that for $0 > \sigma_z > \sigma_{II}$ and $\beta_{II} < 0$, one of the zeros will be in the right half plane - a nonminimum phase effect. As σ_z approaches σ_{II} , the contribution of the phugoid to the adverse motion becomes larger and the phugoid portion is exclusively responsible for that kind of behavior when $\sigma_z \leq \sigma_{II}$.

As an additional interesting fact, there is no pair of conjugate complex zeros for

$$\beta_I > \frac{\omega_{oII}^2}{4\sigma_I(\sigma_{II} - \sigma_I)} = \beta'_I \quad (32)$$

or

$$\beta_{II} < - \frac{\omega_{oI}^2}{4\sigma_{II}(\sigma_{II} - \sigma_I)} = \beta'_{II} \quad (33)$$

For those values of either β_I or β_{II} , the corresponding loci circles degenerate into a point on the real axis at double σ_1 or σ_2 , respectively. On the other hand, the ranges in β_I combined with certain values of β_{II} are also limited for certain regions of zero locations. For instance, in the example shown in figure 10, there is no $\beta_I > 0$ together with $\beta_{II} = 1.0$, that will form a pair of conjugate complex zeros.

For values of β_I and β_{II} , as indicated in equations (32) and (33), the corresponding points in the σ_z, ω_z plane, each representing a pair of real zeros, lie on hyperbolas that do not reach the σ_z axis. As their conjugate axis is the σ_z axis itself, the point closest to the σ_z axis is given by the hyperbola vertex (fig. 10). With growing β_I and decreasing β_{II} , the vertices of the corresponding hyperbolas are moving from the σ_z axis to a maximum displacement, where their vertices lie on the intersection with all other hyperbolas belonging to that loci family. The β_I and β_{II} values for that case are

$$\beta_I^* = 2\beta'_I, \quad \beta_{II}^* = 2\beta'_{II} \quad (34)$$

It appears from figure 10 that there is only a small region of zero pair locations in which significant contribution of the short-period modal response will be expected or equivalently where β_{II} is greater than 1.0. For these values, conjugate complex zeros must be located within the corresponding circle shown close to the origin. For pairs of negative real zeros the same is possible only in the narrow region between the 45°-line from the origin and the hyperbola extending the $\beta_{II} = 1.0$ circle in the lower half plane.

Considering the case of single zeros in terms of β_I and β_{II} , the definition of β' and β^* is also very useful. A single zero located at points

where for pairs of zeros, $\beta_I = \beta_I'$ and $\beta_{II} = \beta_{II}'$ (fig. 10), will have the same values for β_I and β_{II} . This becomes evident, if the relations for β_I , β_{II} , and β_{III} are developed for the single-zero case from

$$N(s) = 1 + b_1s = \beta_I(1 + \alpha_{1II}s + \alpha_{2II}s^2) + \beta_{II}(1 + \alpha_{1I}s + \alpha_{2I}s^2) + \beta_{III} \quad (35)$$

such that, with $\sigma_z = -1/b_1$,

$$\left. \begin{aligned} \beta_I &= -\frac{\sigma_I}{\sigma_z} \cdot \beta_I^* \\ \beta_{II} &= -\frac{\sigma_{II}}{\sigma_z} \cdot \beta_{II}^* \\ \beta_{III} &= 1 - (\beta_I + \beta_{II}) \end{aligned} \right\} \quad (36)$$

These relations are very simple and do not require any plotting, as should be expected for the case of a single zero.

Three Zeros

The extension to the case of three zeros requires a slight augmentation of the transfer function expansion used in the last sections. An additional β coefficient must be generated to keep the main structure of the expansion and ensure the validity of the approximation that was made. The augmented expansion becomes

$$\begin{aligned} F(s) &= \frac{1 + b_1s + b_2s^2 + b_3s^3}{1 + a_1s + a_2s^2 + a_3s^3 + a_4s^4} \\ &= \frac{\beta_I}{D_I(s)} + \frac{\beta_{II} + s\beta_{II}''}{D_{II}(s)} + \frac{\beta_{III}}{D(s)} \end{aligned} \quad (37)$$

Following the approximation made earlier,

$$F(s) \approx \frac{\beta_I + \beta_{III}}{D_I(s)} + \frac{\beta_{II} + s\beta_{II}''}{D_{II}(s)} \quad (38)$$

This essentially means that, in contrast to the case of two zeros, only the short-period portion of the response will be modified by its derivative due to

the gain β_{III}'' . Neglecting the contribution of the β_{III}'' expansion term itself in the first place, β_I , β_{II} , and β_{III} are all implicit functions of the third zero, and can be treated the same way in the case of a pair of zeros. The β plots developed for the case of a pair of zeros (fig. 10) can be used for the interpretation of transfer functions with three zeros. In fact, β_I through β_{III} can be described by the relations of equation (28) by replacing ω_{OZ} and σ_Z by ω'_{OZ} and σ'_Z , as follows:

$$\left. \begin{aligned} \beta_I &= \frac{\omega_{OI}^2}{\omega'_{OZ}{}^2} \cdot \frac{\sigma'_Z - \sigma_I}{\sigma_{II} - \sigma_I} \\ \beta_{II} &= - \frac{\omega_{OI}^2}{\omega'_{OZ}{}^2} \cdot \frac{\sigma'_Z - \sigma_{II}}{\sigma_{II} - \sigma_I} \\ \beta_{III} &= 1 - (\beta_I + \beta_{II}) \end{aligned} \right\} \quad (39)$$

with the interrelations between ω'_{OZ} , σ'_Z , and ω_{OZ} , σ_Z , and z_3 , the real third zero,¹ given by:

$$\left. \begin{aligned} \frac{2\sigma'_Z}{\omega'_{OZ}{}^2} &= \frac{1}{z_3\omega_{OZ}^2} (2z_3\sigma_Z + \omega_{OZ}^2 - \omega_{OI}^2) \\ \frac{1}{\omega'_{OZ}{}^2} &= \frac{1}{z_3\omega_{OZ}^2} (z_3 + 2\sigma_Z - 2\sigma_I) \end{aligned} \right\} \quad (40)$$

which leads to the equation for the equivalent pair of zeros $z'_{1,2}$:

$$z'_{1,2} = -A \pm \sqrt{A^2 - B^2} \quad (41)$$

where

$$\begin{aligned} A &= \frac{-2z_3\sigma_Z - \omega_{OZ}^2 + \omega_{OI}^2}{2(-z_3 - 2\sigma_Z + 2\sigma_I)} \\ B^2 &= \frac{z_3\omega_{OZ}^2}{z_3 + 2\sigma_Z - 2\sigma_I} \end{aligned}$$

A root locus plot for $z'_{1,2}$ as function of the parameter $Q = z_3(\sigma_Z - \sigma_I)/2$ can be derived (fig. 11). The root locus equation becomes:

¹The third zero z_3 is chosen as either the single real zero or arbitrarily from the three real zeros.

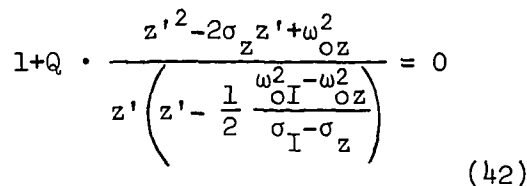


Figure 11.- Equivalent pairs of zeros
for use of β_I and β_{II} plot in
the case of three zeros.

Although the methods described in the previous sections can be useful for analysis purposes in the frequency domain, they are used in this section for the evaluation of an optimal design.

Linear optimal control theory optimizes the time response of a given system with respect to a quadratic performance criterion, weighting the states as well as the controls. In the case of the longitudinal motion of aircraft, the relation between the requirements on the time response and the weighting coefficients of the performance index is not strictly clear. Thus, only a hypothetical guess of the weighting coefficients can be made that has to be proved satisfactory by examining the actual performance of the corresponding time response and control laws. The dependence of the time response on the performance index may be studied effectively by the procedure outlined above.

Here we study this dependence for the aircraft longitudinal control system.

The system investigated as an example is the regulator-controlled longitudinal motion of a conventional subsonic-type transport aircraft during landing approach (fig. 12). It can be described by the state equations:

where x is the state error vector

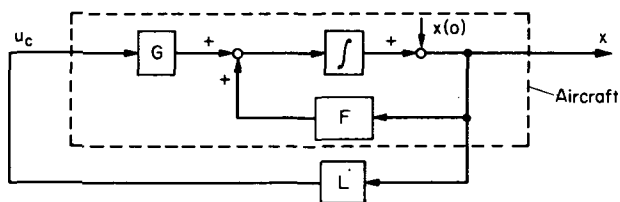


Figure 12.- System block diagram.

$$x = \begin{bmatrix} \dot{\theta} \\ \theta \\ \gamma \\ \hat{u} \end{bmatrix}$$

u_c is the control vector

$$u_c = \begin{bmatrix} \delta_e \\ \delta_T \\ \delta_F \end{bmatrix}$$

and F , G , and L are matrices with constant elements. The F and G matrices are given as aircraft parameters, while the elements of L have to be determined as the regulator feedback gains. This type of aircraft control is adequate when the flight path is prescribed by air-traffic control as a straight line, such that an equilibrium state can be assigned to it. Thus the task of the controller is to drive the errors of the states, caused by initial displacements $x(0)$, to zero. Performing this task in the described setup, the Regulator could consist of an automatic controller, or both automatic and manual in parallel. The total gains would be equal to L . If the gain portion due to manual control is kept inside the allowable range for pilot gains, there might be no further dynamical equalization required from the pilot to minimize the performance criterion chosen for the system.

Performance Criterion

To follow a given straight flight path corresponding to a certain equilibrium state, the task of the controller can be reduced to that of a regulator, mainly driving the magnitude of the error in the speed vector to zero.

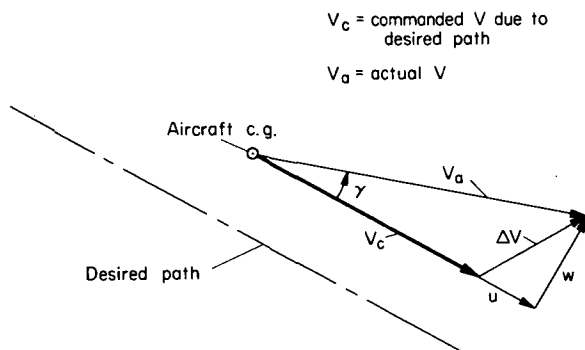


Figure 13.- Definition of the speed notations.

The reduction of the magnitude of the speed error vector ΔV (fig. 13) is of primary interest with respect to the overall behavior, irrespective of its direction. Thus, there is only one independent quantity describing the performance of the system output instead of the original four given by the components of the error state vector.¹ This single error has to be weighted against the allowable effort in the controls with all controls having the same weighting. The magnitude of the error in the speed vector V_C can be expressed in terms of error state components, as shown in figure 13:

¹For simplicity, the penalty of the error in height is ignored. Computations have shown that this does not result in a big difference in the resulting optimal control law.

$$|\Delta V|^2 = u^2 + w^2 \quad (44)$$

$$\text{With } w = (u + |V_C|) \cdot \tan \gamma \quad (45)$$

and assuming small γ as well as small u compared to $|V_C|$, equation (44) becomes

$$|\Delta V|^2 = u^2 + (|V_C| \cdot \gamma)^2 \quad (46)$$

Thus, the minimum of the quadratic performance index

$$T = \frac{1}{2} \int_0^\infty [q(u^2 + |V_C|^2 \gamma^2) + r(\delta_e^2 + \delta_T^2 + \delta_F^2)] dt \quad (47)$$

defines an optimal regulator control law

$$u_c = Lx \quad (48)$$

with respect to the value chosen for the weighting ratio q/r , where $q > 0$ and $r > 0$ (refs. 6, 7).

Discussion of Resulting Optimal Transfer Functions

Now, suppose that L has been computed in the standard way. Taking the Laplace transform of equation (43) with initial condition $x(0)$, the response is given by a 4×4 matrix H ,

$$x(s) = [Is - (F + GL)]^{-1} x(0) = H \cdot x(0) \quad (49)$$

whose elements can be considered as transfer functions with the error state components as outputs and the initial conditions as δ -function inputs. The elements of H are easily derived from equation (49):

$$h_{ij} = \frac{x_i(s)}{x_j(0)} = \frac{\det[Is - (F + GL)]_{i,j}}{\det[Is - (F + GL)]} \quad (50)$$

where $[.]_{i,j}$ is equal to $[.]$ in all elements except the i th column, which is replaced by the j th column of the 4×4 unity matrix.

It will be shown that the main changes in the character of the time response take place in the range of q/r between 0 and 2.0. Beyond that, the control effort becomes higher and higher, with no comparable significant changes in the time responses of the error states. For $q/r = 2.0$, the phugoid poles are still conjugate complex, so that only the first described method is used.

The aircraft system has three controls and three states of particular interest, resulting in nine separate responses or transfer functions to be considered. Time histories of these states and controls for q/r values of 0 and 2.0 are shown in figures 14 through 17. The quantity $\dot{\gamma}$ is also plotted to show the load factor. There are considerable changes in the time histories with increasing q/r . Figures 18 through 20 show the loci required for the application of this method. Figure 18 shows the poles and zeros for all the various transfer functions for $q/r = 0$ and 2.0. In addition, the phase loci are drawn for $q/r = 2.0$ for both $p_1 = p_1$ and $p_1 = p_3$. Figure 19 shows the poles and zeros, but also includes the magnitude loci associated with the

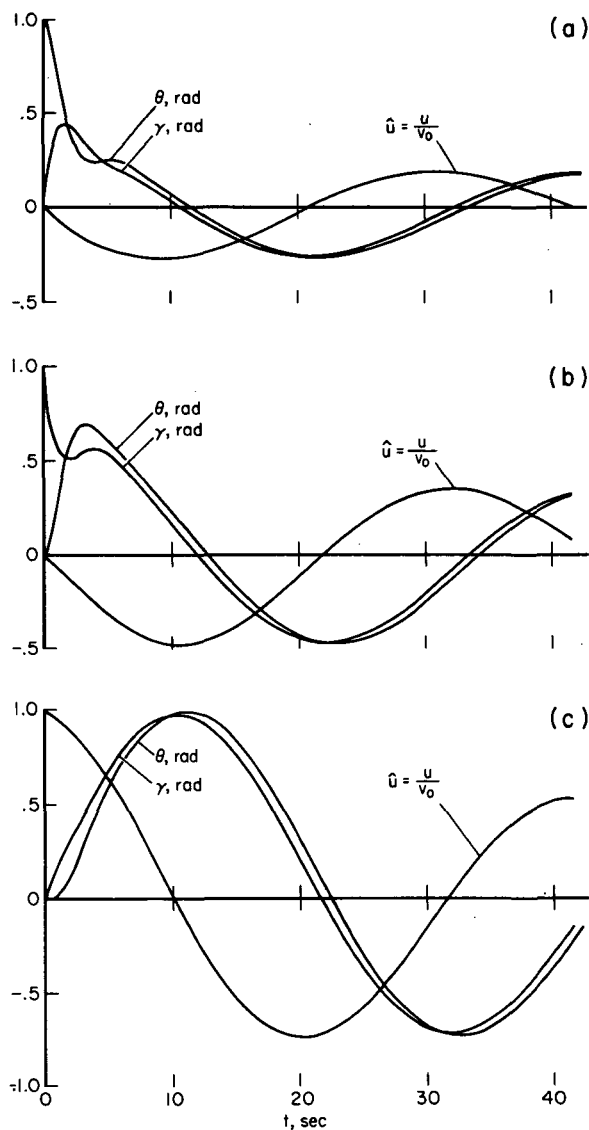


Figure 14.- Transient responses due to initial conditions of original system ($q/r = 0$); (a) in θ , (b) in γ , (c) in \hat{u} .

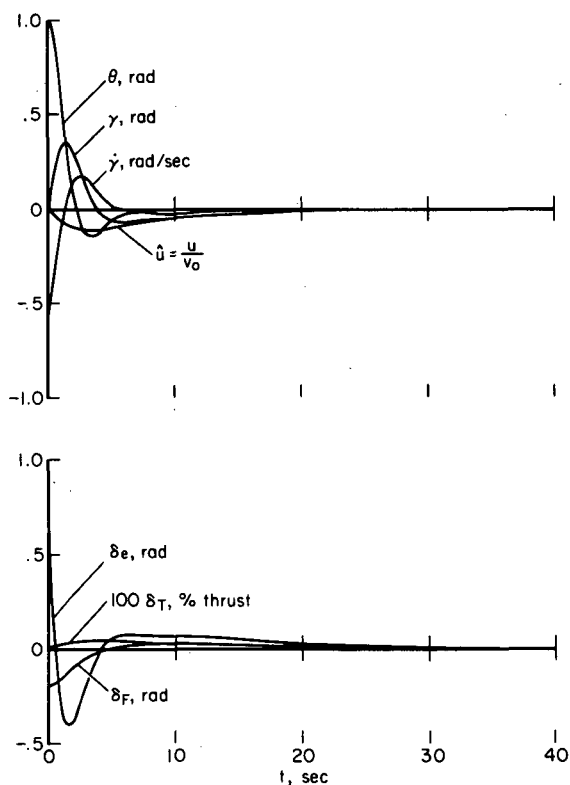


Figure 15.- Transient responses of the state and control variables of the regulated system ($q/r = 2.0$) due to initial condition in θ .

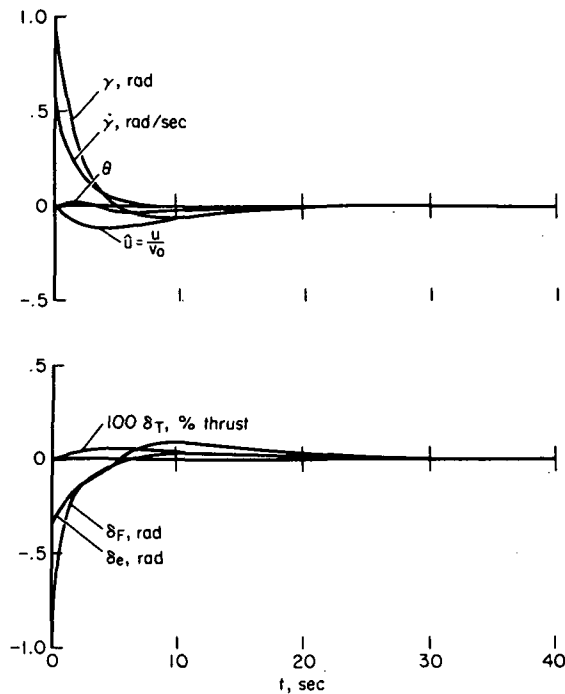


Figure 16.- Transient responses of the state and control variables of the regulated system ($q/r = 2.0$) due to initial condition in γ .

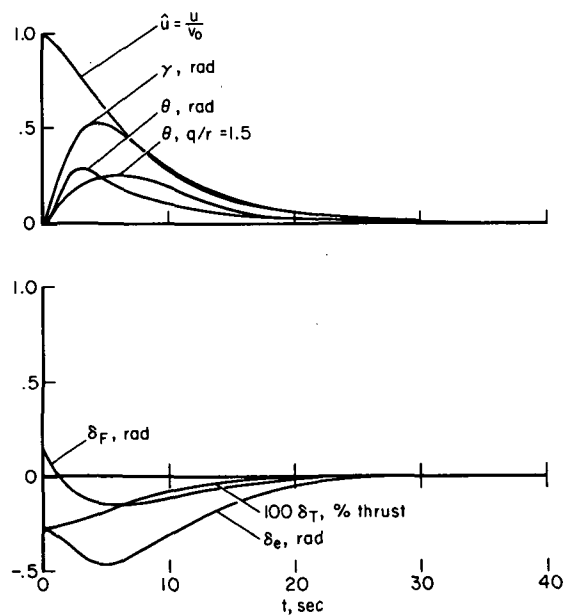


Figure 17.- Transient responses of the state and control variables of the regulated system ($q/r = 2.0$) due to initial condition in \hat{u} .

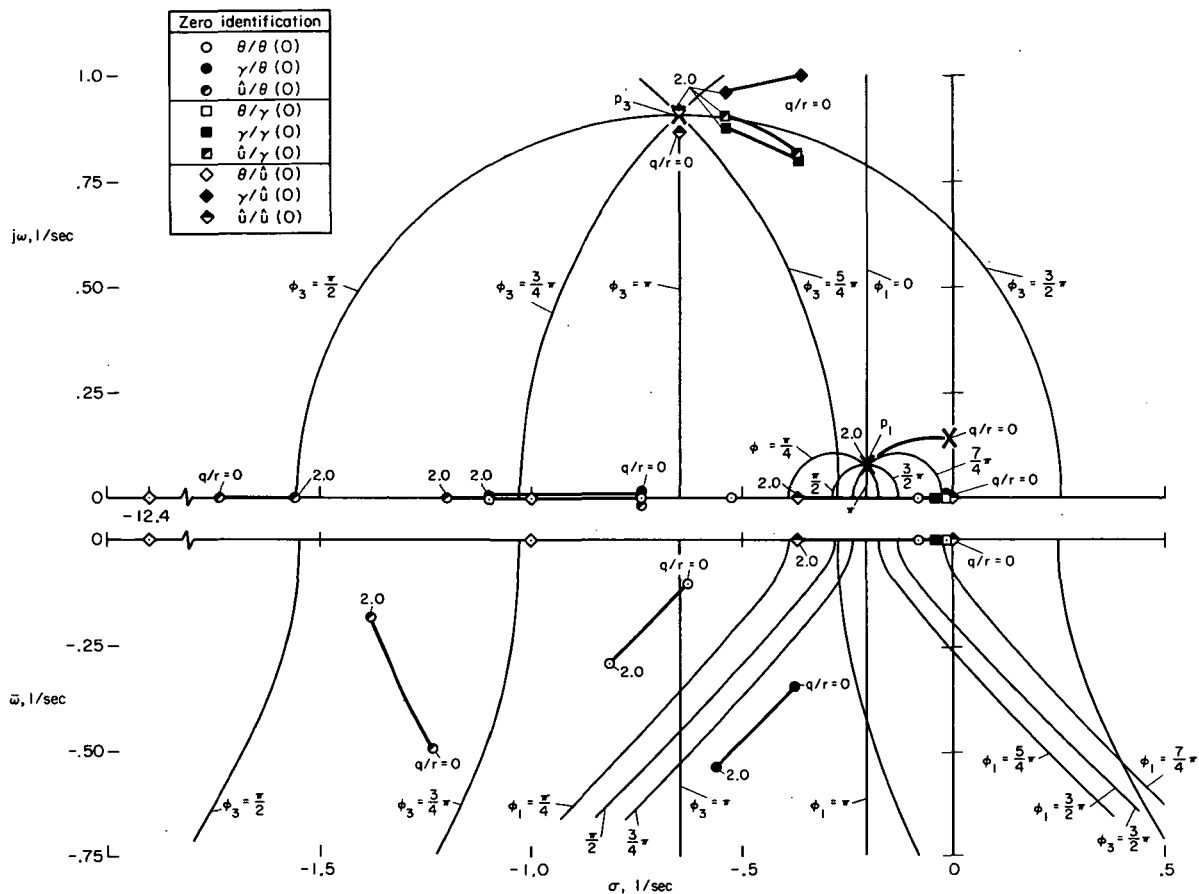


Figure 18.- Phase loci for longitudinal response residues.

phugoid residues again for $q/r = 2.0$. Figure 20 is similar to 19, with the magnitude loci drawn for the short-period residues. A number of the roots change only slightly with q/r , and some change very similarly for several transfer functions.

In deriving the time vectors (or residues) by means of figures 18 through 20, we use as an example the transfer function $\gamma/\hat{u}(0)$ for $q/r = 0$, which has a second-order numerator and a fourth-order denominator. As shown earlier, the time vector of the short-period mode γ_3 , for instance, will be determined from the residue C_3 (at p_3), which in turn consists of the product of C_{3z} and C_{3p} (eq. (6)). The phase of $C_{3z}\varphi_{3z}$ is determined from p_3 and z_1 as indicated in figure 21(a), which shows the appropriate poles and zeros extracted from figure 18. The conjugate zero z_2 does not need to be included for this plot. The quantity $|C_{3z}| = c_1 \cdot c_2$ is obtained from figure 21(b), which is abstracted from figure 20, using p_3 and z_1, z_2 . These quantities, φ_{3z} and $|C_{3z}|$, are then combined and plotted in figure 22 as C_{3z} . A similar process, shown in figure 21(c) and (d), is used to derive $p_3 C_{3p}$.¹ The phase

¹Note that $p_3 C_{3p}$ or $p_3 C_3$ is used due to the δ -function input $\hat{u}(0)$ instead of a step input.

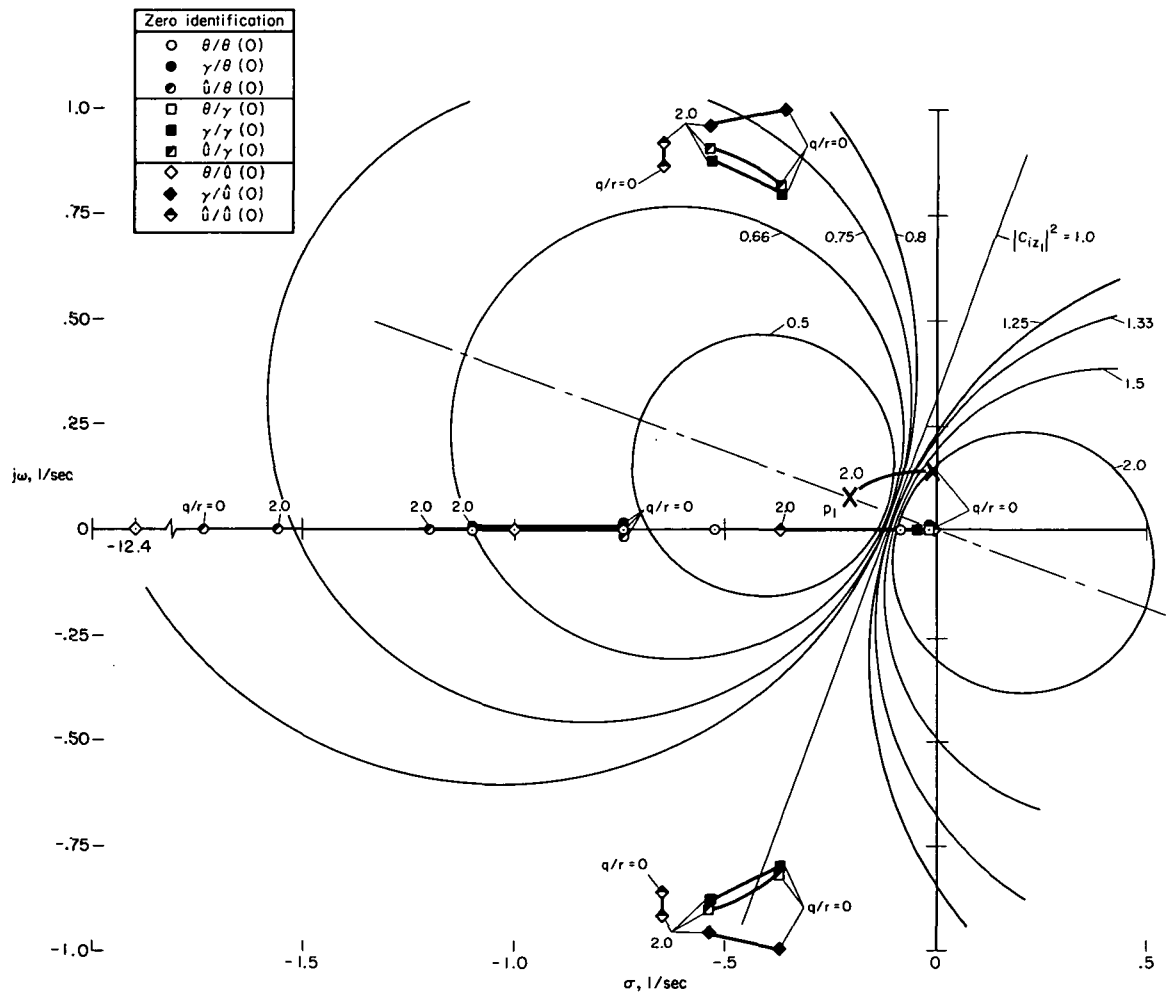


Figure 19.- Magnitude loci for longitudinal response phugoid residues.

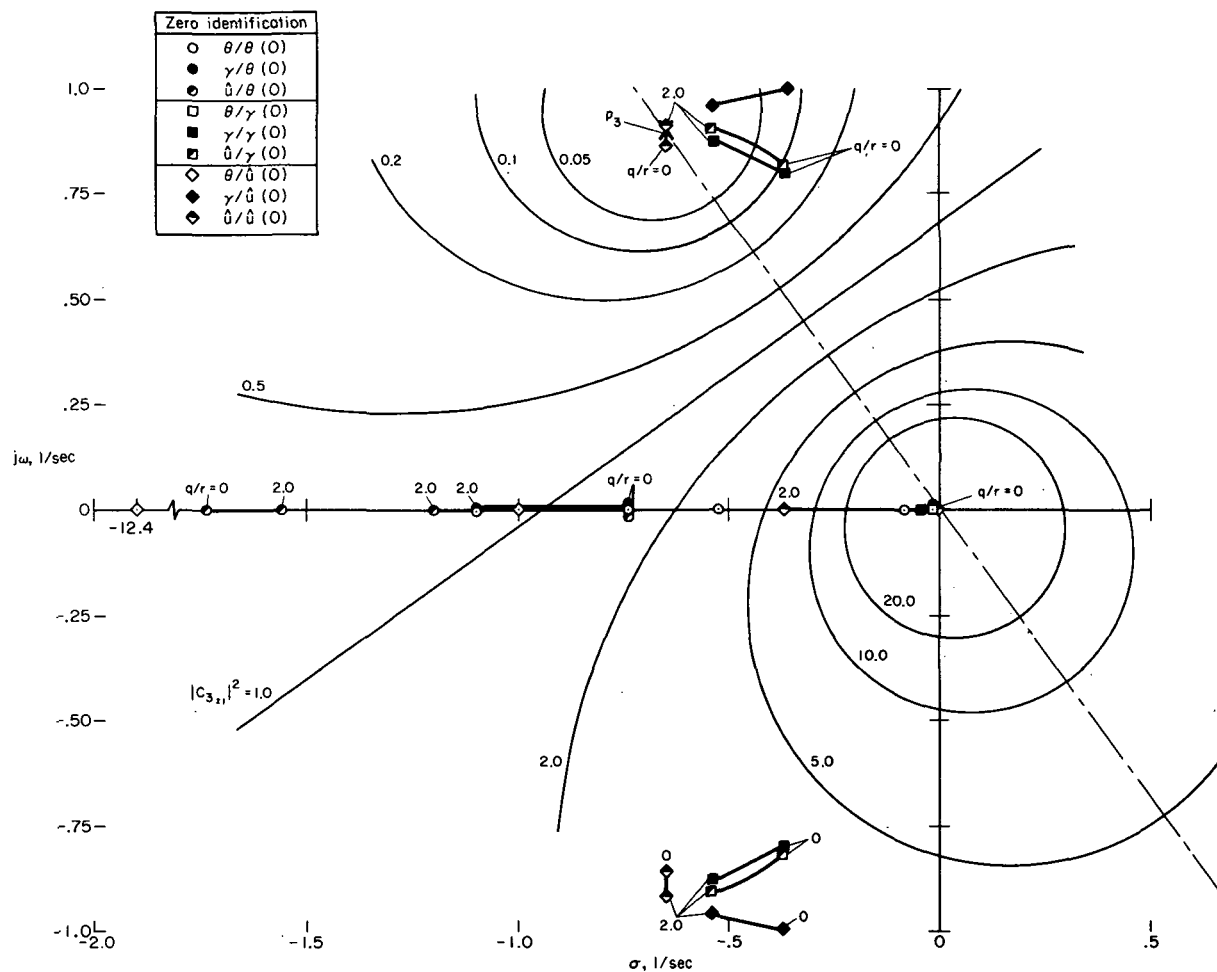


Figure 20.- Magnitude loci for longitudinal response short-period residues.

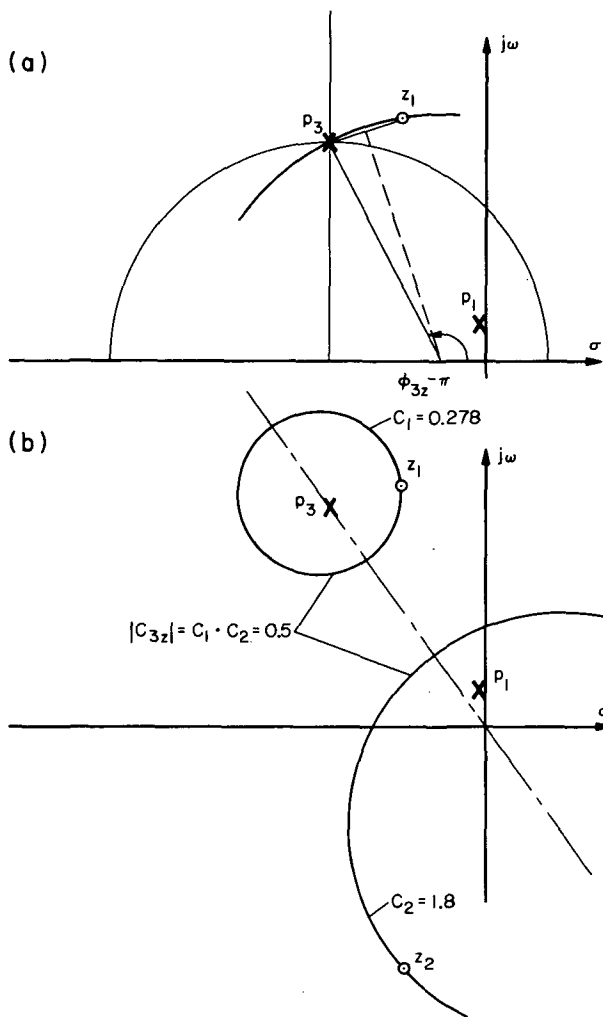


Figure 21.- (a) Phase locus and (b) magnitude loci for C_{3z} of the transfer function $\gamma/\hat{u}(0)$; $q/r = 0$.

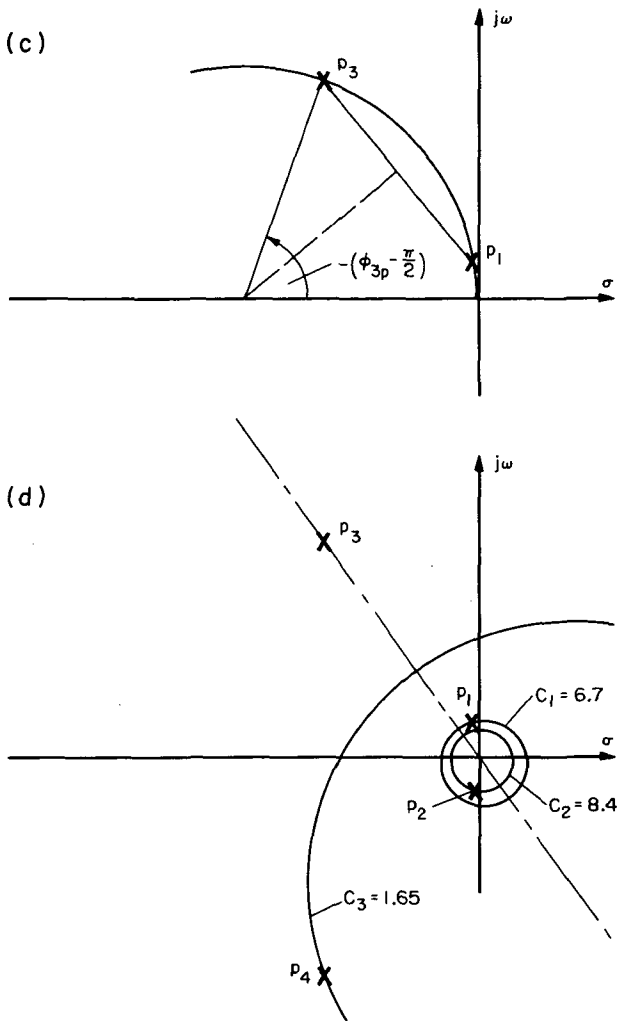


Figure 21.- (c) Phase locus and (d) magnitude loci for $p_3 C_{3p}$ of the transfer function $\gamma/\hat{u}(0)$; $q/r = 0$.

of $p_3 C_{3p}$, denoted as ϕ_{3p} in figure 22, is obtained by using p_1 instead of z_1 (fig. 21(c)). The angle of $\pi/2$ must be added due to the contribution of p_4 , the conjugate pole of p_3 , and the sign has to be changed (eq. (22)). The magnitude of $p_3 C_{3p}$ is given by the product of $|p_3|$ and $|C_{3p}| = 1/(c_1 \cdot c_2 \cdot c_3)$ which in turn is obtained from figure 21(d), using all poles p_1 through p_4 . Again, phase and magnitude are combined to plot $p_3 C_3$ in figure 22.

The quantities C_{1z} and $p_1 C_{1p}$ for the time vector $\bar{\gamma}_1$ are obtained similarly using p_1 instead of p_3 .

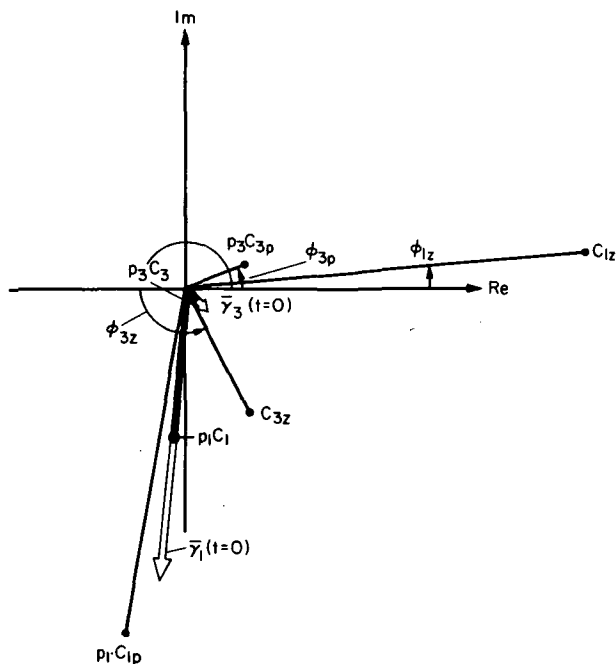


Figure 22.- The time vector construction from the residue quantities.

With p_1C_1 (p_3C_3) determined, the actual time vectors $\bar{\gamma}_1$ ($\bar{\gamma}_3$) at $t = 0$ (fig. 22) are two times p_1C_1 (p_3C_3). This same process is used to determine all the time vectors shown in the subsequent figures.

In the following, the time responses of the state components θ , γ , and \hat{u} due to initial conditions in the same states are discussed separately, referring to figures 18 through 20. The time histories of the controls could be discussed in the same way.

$\theta/\theta(0)$ —The attitude is the main control variable in the longitudinal motion of aircraft. The requirement is to reduce the initial attitude error in a rapid, well-damped manner to zero. Without any control ($q/r = 0$), the time response behavior of the attitude due to initial displacement shows large overshoot and very little damping of the low-frequency mode, the phugoid (see

fig. 14), whereas for optimal control with respect to the chosen performance index and increasing q/r the desired behavior can be achieved, as shown in figure 15.

There are three real zeros but the overall tendencies in the change of the time response due to increasing q/r are caused mainly by the shift of the phugoid poles to higher damping and lower frequency. This shift will affect all time responses, because the poles are the same for all transfer functions for a given q/r . Therefore, for all transfer functions, there will be the tendency for C_1 , the phugoid mode residue or half of the time vector $\bar{\theta}_1$ at $t = 0$, to shift its phase into the imaginary axis (fig. 23), as the phugoid poles approach the real axis of the root plane. (A similar effect will occur for other transfer functions.) Because there is almost no change in the short-period poles with increasing q/r (also common for all transfer functions), the location of the zeros with respect to the phugoid poles determines how far the phase shift in $\bar{\theta}_1$ will go and what change in magnitude of that residue will occur.

As can be seen from figures 18 and 19, $1/C_p$ increases mainly in magnitude and C_{1z} in phase with higher values of q/r . The phase shift is chiefly due to the shift of the phugoid poles in relation to the zero that is closest to the origin. The change in the phase contribution of that zero amounts to almost $\pi/2$ in the range of q/r from 0 through 2.0. Although there are also changes taking place due to the zero farthest from the origin,

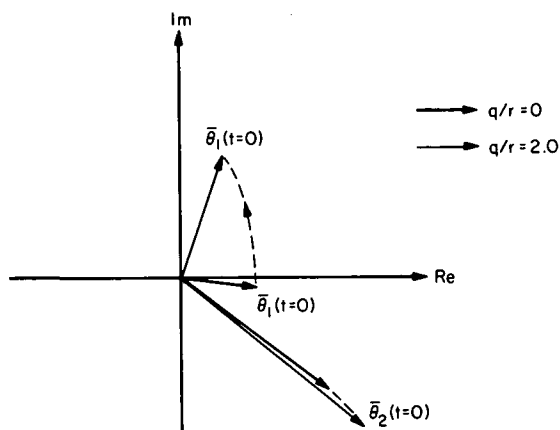


Figure 23.- Time vectors for the transient response of $\theta/\theta(0)$ with $q/r = 0$ and 2.0 .

axis and the high phugoid damping. On the other hand, choosing a well-damped location for the phugoid and fixing one of the zeros, a given range of the C_{1z} -phase loci can be assigned to the location of the remaining zeros to achieve the desired behavior.

$\gamma/\theta(0)$ and $\hat{u}/\theta(0)$ —Although suppression (decoupling) of the responses in the flight-path angle γ and the speed \hat{u} due to initial displacement in the attitude might be desired, a certain amount of temporary displacement in both γ and u is required to avoid excessive costs in the controls. That is, a very effective lift control would be needed, with simultaneous speed control, to compensate for the initial lift change due to $\theta(0)$.

As can be seen from figures 14 and 15, at least some improvement is made, particularly in the speed response, by minimizing the previously described performance index. Again, the main reason is the shift of the phugoid poles toward the real axis of the root plane, which leads to high damping and fixes the phase of $\bar{\gamma}_1$ and \bar{u}_1 close to the imaginary axis for both the γ and u response (fig. 24).

To determine the actual phase of $\bar{\gamma}_1$ and $\bar{\gamma}_3$, one can make use of the results obtained by examining the zeros for $\theta/\theta(0)$ that are located in close vicinity to both existing zeros of $\gamma/\theta(0)$. Therefore, the phase difference between the residues of the responses in attitude and flight-path angle is caused by the phase effect of the medium zero of $\theta/\theta(0)$. Because this zero does not change with q/r , the phase difference between the residues of $\theta/\theta(0)$ and $\gamma/\theta(0)$ is also independent of q/r .

Response due to $\gamma(0) \neq 0$ —As shown in figure 17, the response to an initial displacement in the flight-path angle is good for the case of $q/r = 2.0$. The flight-path angle itself settles down to small values very soon without evident oscillations, and the coupled motion in pitch and speed is almost negligible. Also, the load factor due to the rate of change of the flight-path

that zero has little effect on either the phase or magnitude of C_{1z} . Thus, the tendency in C_{1z} is simply obtained by watching only the relative location of the phugoid poles and the zero close to the origin. With respect to $\bar{\theta}_3$ or C_3 (the residue of the short-period mode), the change in C_{3p} is due to the phugoid shift and the change in C_{3z} is due to the shift of the zero far from the origin. These are opposing changes, such that the total change in C_3 is rather small.

The conclusion is that the effect of the phugoid mode is considerably reduced by the large phase shift in $\bar{\theta}_1$ toward the imaginary

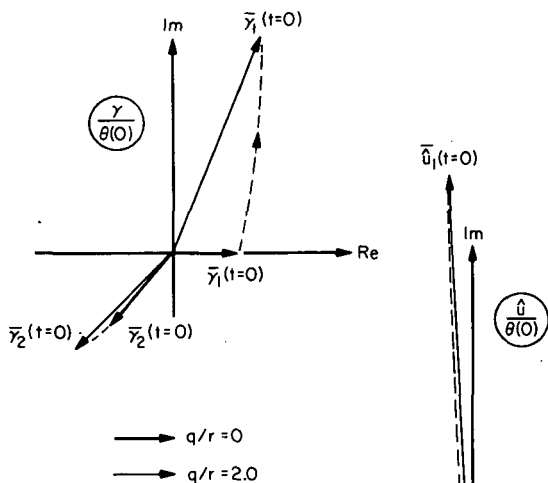


Figure 24.- Time vectors for transient responses of $\gamma/\theta(0)$ and $\dot{u}/\theta(0)$ with $q/r = 0$ and 2.0 .

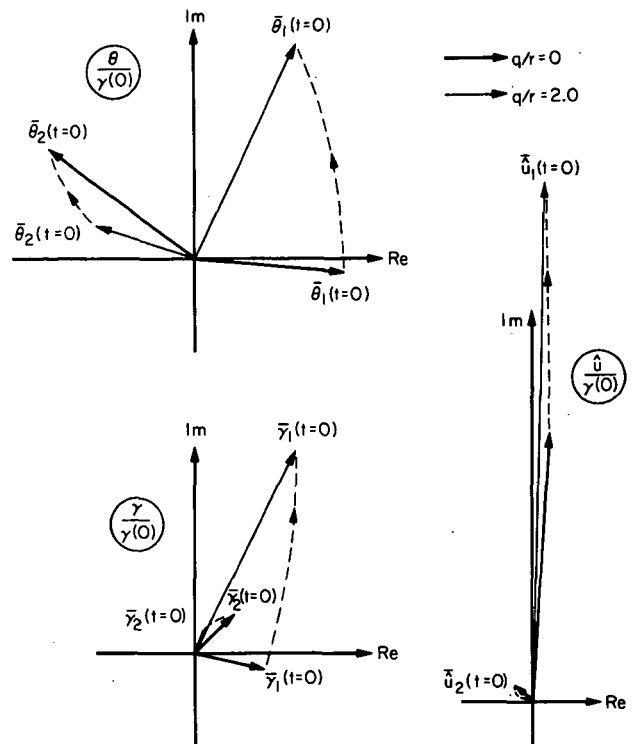


Figure 25.- Time vectors for transient responses of $\theta/\gamma(0)$, $\gamma/\gamma(0)$, and $\dot{u}/\gamma(0)$ with $q/r = 0$ and 2.0 .

angle does not appear to yield too high values. The plot of the time vectors of the response in θ is shown as part of figure 25. The short-period vector is phased with respect to the phugoid vector such that while rotating and before decaying, their real parts almost cancel, resulting in low coupling between θ and $\gamma(0)$. This configuration of the vectors thus shows a highly decoupled transfer function, and a similar configuration would be desirable for other transfer functions. Comparing this configuration with that of $\gamma/\theta(0)$ (fig. 24), for instance, a smaller phase shift for $\bar{\gamma}_3$ would produce less coupling. This could be obtained by locating the zeros on the corresponding phase locus curve. These phase values are easily obtained from figure 25 since only one zero is involved in the $\theta/\gamma(0)$ response.

The pair of conjugate complex zeros of $\gamma/\gamma(0)$ is almost coincident with the zeros of $\dot{u}/\gamma(0)$ and $\gamma/\dot{u}(0)$. Thus, the difference in the behavior of these later two in comparison to $\gamma/\gamma(0)$ is given by the real zero of $\gamma/\gamma(0)$, which again is almost independent of q/r . Observing the magnitude ratio between the residues $\bar{\gamma}_1$ and $\bar{\gamma}_3$ in figure 25, one can see that increasing q/r results in a high gain for $\bar{\gamma}_1$ relative to $\bar{\gamma}_3$, which gives a very slow long-term dynamic behavior.

Response due to $u(0) \neq 0$ —According to figure 17, the direct response in the speed again appears satisfactory for moderate values of q/r . There is no fast response required for the speed so long as the initial error returns to

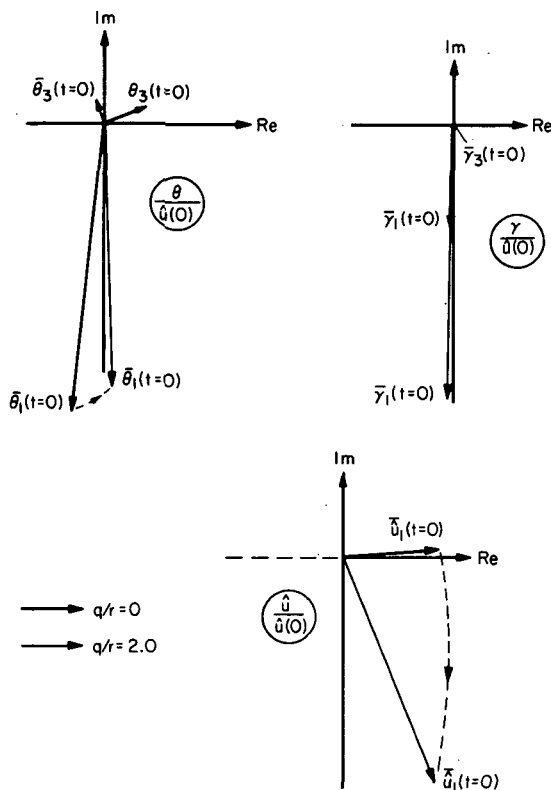


Figure 26.- Time vectors for transient responses of $\theta/\hat{u}(0)$, $\gamma/\hat{u}(0)$, and $\hat{u}/\hat{u}(0)$ with $q/r = 0$ and 2.0 .

the θ response, a q/r less than 2.0 might do more to minimize the peak in the θ -time history. It can be deduced from figure 26, by considering the motion of the time vector $\bar{\theta}_1$ for values of t greater than 0 , that a smaller q/r will reduce slightly the peak in θ , but increase the long-period response. This result is shown on figure 17. Note that the other responses will also change, which must be considered.

Abstract of the results—The simple performance index, weighting only the error state in the flight-path angle and the airspeed against the controls, using one parameter for each, provides in the main satisfactory results. For $q/r \approx 2$ most of the transient responses are satisfactory, but two of them are only marginally acceptable.

As shown in figure 17, improvement in those time responses would require a different performance index. Since the error in γ is relatively more annoying than that in \hat{u} , a lower weighting of \hat{u} compared to γ would be desirable. Although this improvement is not carried out in this study, it would require at least one additional weighting parameter, and two parameter patterns of root loci drawn coincident with the optimal results for different levels of weighting in both parameters. The methods presented here could

zero in a monotonic but not too sluggish manner. This is achieved by the high damped, practically pure phugoid motion (fig. 26) due to the pair of conjugate complex zeros very close to the short-period poles (figs. 18-20). On the other hand, the coupling of the velocity to flight-path angle and the pitch angle is fairly large, giving somewhat unsatisfactory responses. Although there is a considerable improvement in the overall dynamic behavior for $q/r = 2.0$, compared to the case without regulator control ($q/r = 0$), the gains K of the transfer functions, especially of $\gamma/\hat{u}(0)$, are still high.

The effect of the zero of $\theta/\hat{u}(0)$, which has a large decrease in value as q/r varies from 0 to 2.0 (figs. 18-20), is to cause a large phase shift of the short-period time vector (residue) (fig. 26). Also, the phugoid vector moves across the imaginary axis before merging with it for $q/r > 2$. Considering that both the phase shift of the residues and the high damping of the phugoid are responsible for the improvement in

still be used to aid in the determination of the elements of q and r that would give the desired time responses.

CONCLUDING REMARKS

Two methods are described that provide geometrical insight into the effect of the zeros of transfer functions as well as poles, on the corresponding time histories. The main advantage of these methods lies in their simplicity compared to other available procedures. They differ in their limitations, which are essentially those on the possible regions of the system pole locations.

The first method, which uses the geometrical interpretation of the residues, loses much of its power if the poles become real, while the other method is restricted to certain separations of the poles, regardless of whether they are real or conjugate complex. In almost all cases, at least one of the methods can be used effectively. This is especially true when applied to aircraft transfer functions, as shown in the discussion of optimal regulator control for the longitudinal motion.

Both methods are developed for the type of transfer function with four poles and two zeros along with an extension to three zeros. Further extension to allow for additional poles and zeros is possible, but will be very complicated.

The main use of the methods should be for analysis purposes. To interpret the effect of zero changes, the poles are assumed fixed, and changes in the poles are discussed with the zeros unchanged. Thus, only if something is known about the sensitivities of the pole locations due to zero changes and vice versa (ref. 8) can the methods be helpful in the synthesis process.

The methods are applied to a study of the longitudinal motion regulator of an aircraft. The performance of this regulator system is studied using a particular performance criterion with minor variations. This criterion minimizes the state errors effectively using only reasonable control effort.

Ames Research Center
National Aeronautics and Space Administration
Moffett Field, Calif. 94035, Feb. 3, 1972

APPENDIX A

EQUATION FOR THE MAGNITUDE LOCI OF C_{iz} (TWO ZEROS)

The exact equation for the magnitude loci of C_{iz} is simply derived from equation (14) by substituting in the expressions for c_1 and c_2 the values of the conjugate complex zeros, $z_{1,2} = \sigma_z \pm j\omega_z$:

$$|C_{iz}| = c_1 \cdot c_2$$

$$= \frac{[(\sigma_z - \sigma_i)^2 + (\omega_z - \omega_i)^2][(\sigma_z - \sigma_i)^2 + (\omega_z + \omega_i)^2]}{(\sigma_z^2 + \omega_z^2)} \quad (A1)$$

The coordinates of the pair of conjugate zeros generate fourth-order terms, such that the loci of $|C_{iz}|$ in the root plane become higher order curves. For the special case of $|C_{iz}| = 1$, which might be of particular practical interest, equation (A1) reduces to

$$4\sigma_z^3\sigma_i - 2\sigma_z^2(\omega_i^2 + 3\sigma_i^2) + 4\sigma_z\sigma_i |p_i|^2 + 4\sigma_z\omega_z^2\sigma_i - 2\omega_z^2(\sigma_i^2 - \omega_i^2) - |p_i|^4 = 0 \quad (A2)$$

From this equation, the simple formulas (17) through (20) are derived for specific points on that curve by substituting the specific values for σ_z or ω_z .

For a pair of real zeros, equation (A2) changes to

$$4\sigma_z^3\sigma_i - 2\sigma_z^2(\omega_i^2 + 3\sigma_i^2) + 4\sigma_z\sigma_i |p_i|^2 - 4\sigma_z\bar{\omega}_z^2\sigma_i + 2\bar{\omega}_z^2(\sigma_i^2 - \omega_i^2) - |p_i|^4 = 0 \quad (A3)$$

APPENDIX B

ENCIRCLING OF ZERO REGIONS FOR GIVEN VALUES OF $|C_{iz}|$

Practicable approximations can be made to encircle zero regions for given values of $|C_{iz}|$ rather than to define the exact loci curves. This appendix gives some further details on those approximations.

According to equations (15) and (16) for the loci circles of c_1 and c_2 , figure 27 illustrates the relation between the location of the center points of those circles and the corresponding values of c_1 and c_2 . It shows that on the string of center points of the c_1 circles, the corresponding value of c_1 is

$$c_1^2 = 1 + \frac{1}{n} > 1 \quad (B1)$$

if the center point lies at a distance $n|p_i|$ downward from the origin and

$$c_1^2 = 1 - \frac{1}{n} < 1 \quad (B2)$$

if the center point lies $n|p_i|$ upward from the origin. The centers of the c_2 circles are symmetrical about the real axis with the corresponding c_1 centers.

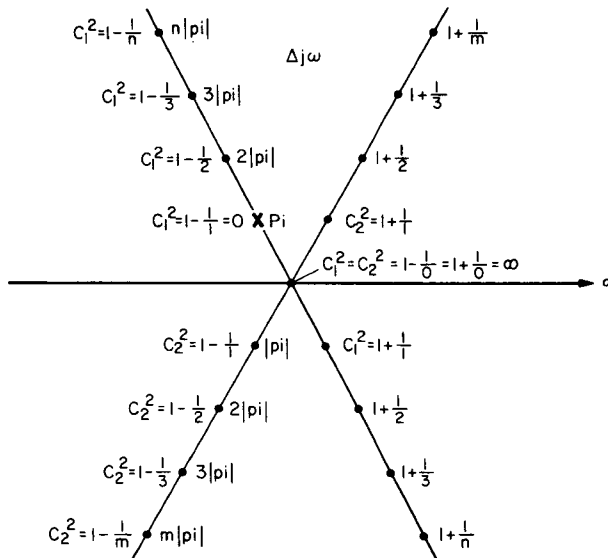


Figure 27.- Relation between values of c_1, c_2 and the location of the loci center points.

In looking for the intersections between c_1 and c_2 circles, it is evident that not all c_1 circles intersect a particular c_2 circle and vice versa. Therefore, considering only the upper half of the root plane for symmetry reasons, there are three regions divided by the straight-line loci of $c_1 = 1$ and $c_2 = 1$ (fig. 7), such that (1) region I includes the intersections possible between the circles of $c_1 < 1$ and $c_2 > 1$, (2) region II includes the intersections possible between the circles of $c_1 > 1$ and $c_2 > 1$, and (3) region III includes the intersections possible between the circles of $c_1 < 1$ and $c_2 < 1$. Encirclement of zero location areas due to certain values of $|C_{iz}|$ is given below for each region.

Region I

If one considers a possible intersection between circles of $c_1 < 1$ and $c_2 > 1$ ($n > 1, m > 0$), the corresponding magnitude of C_{iz} can be written with use of equation (14):

$$|C_{iz}|^2 = \left(1 - \frac{1}{n}\right) \left(1 + \frac{1}{m}\right) \quad (B3)$$

or as an explicit relation for m as a function of n with the parameter $|C_{iz}|$, as illustrated in figure 28,

$$m = \frac{n - 1}{(|C_{iz}|^2 - 1) \cdot n + 1} \quad (B4)$$

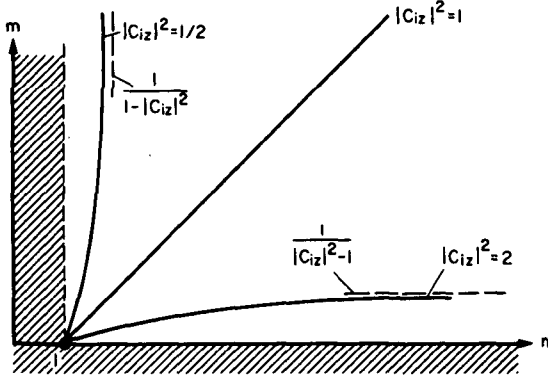


Figure 28.- $|C_{iz}|$ as function of m and n (region I).

The figure shows that for values of $|C_{iz}|$ distinctly smaller (larger) than 1, n (m) does not go over a certain small value. Therefore, the range of circles in c_1 (c_2) corresponding to those $|C_{iz}|$ values is also very limited. Thus, for a certain value of $|C_{iz}|$, the corresponding zeros and all those for smaller (greater) values of $|C_{iz}|$ must lie within the largest possible circle for c_1 (c_2), which is that for $c_1 = |C_{iz}|$ ($c_2 = |C_{iz}|$) (see fig. 7). The range of possible combinations of m and n is not only limited as shown in figure 28; for higher values of n and m , the circles might not intersect such that no locus point for $|C_{iz}|$ can be defined in the root

plane. The upper limit for m and n can be easily described by the following condition, which depends on the location of the pole p_i :

$$|c_1 \pm c_2| = 2 \frac{\omega_i}{|p_i|} \quad (B5)$$

Figure 28 itself is valid for all locations of p_i .

Region II

Within this region the relation between the location of the circles for $c_1 > 1$ and $c_2 > 1$ ($n > 0, m > 0$) and $|C_{iz}|$ itself is

$$|C_{iz}|^2 = \left(1 + \frac{1}{n}\right) \left(1 + \frac{1}{m}\right) > 1 \quad (B6)$$

or

$$m = \frac{n + 1}{(|C_{iz}|^2 - 1)n - 1} \quad (B7)$$

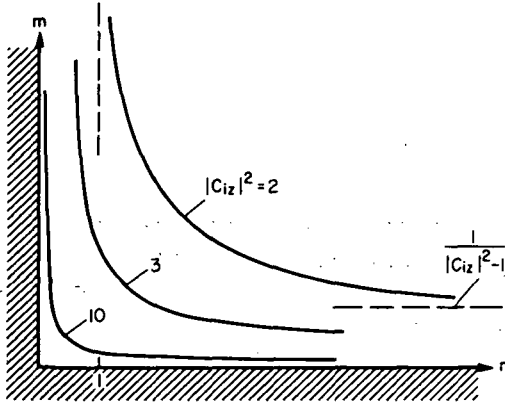


Figure 29.- $|C_{iz}|$ as a function of m and n (region II).

as illustrated in figure 29, which again is independent of the location of the pole p_i . The curves are symmetrical with respect to the straight line $m = n$. Calling $n = n^*$ and $m = m^*$ for the point $m = n$, for each curve it can be stated that for all $n > n^*$ (lower half of the curve) $m < m^*$, and also for all $m > m^*$ (upper half of the curve) $n < n^*$. Therefore, each point on the lower half of the curve represents a zero location that must be inside the c_2 circle given by $m = n^*$ and outside the c_1 circle given by $n = n^*$. A corresponding remark can be given for the upper part of the curve. Thus, the loci for $|C_{iz}|$ can be located in the root plane as: (1) being within the union but outside the intersection of the circles for c_1 and c_2 corresponding to $n = n^*$ and $m = m^*$, respectively, where

$$c_1 = c_2 = |C_{iz}|^{1/2}$$

and (2) intersecting the real axis at the same points, where the circles of (1) are intersecting (fig. 30).

From equation (B6), it follows that

$$n^* = m^* = \frac{1}{|C_{iz}| - 1} \quad (B8)$$

As in region I, there is an upper limit for m and n , dependent on p_i , which is defined as follows:

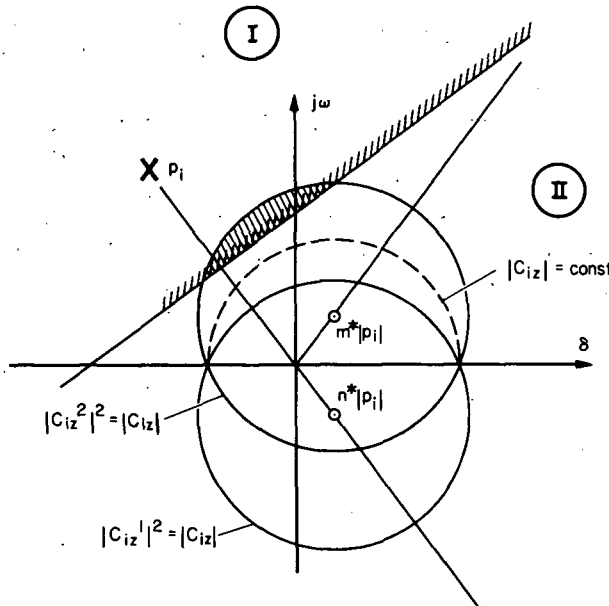


Figure 30.- Approximate region for $|C_{iz}|$ locus (region II).

$$|c_1 - c_2| = 2 \frac{\omega_i}{|p_i|} \quad (B9)$$

This condition can be used, of course, for further encircling.

Region III

Analogous to the approximation applied to region II, for region III, with $c_1 < 1$ and $c_2 < 1$ ($n > 1$, $m > 1$),

$$|c_{iz}|^2 = \left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{m}\right) < 1 \quad (B10)$$

or

$$m = \frac{1 - n}{(|c_{iz}|^2 - 1)n + 1} \quad (B11)$$

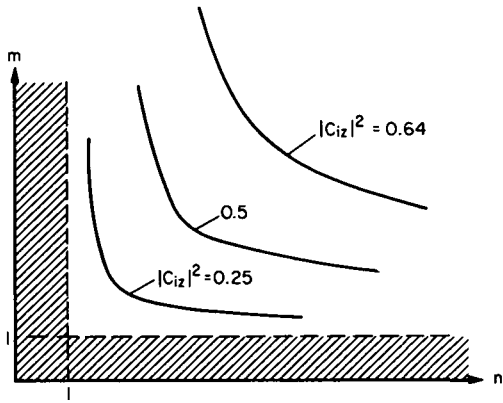


Figure 31.- $|c_{iz}|$ as a function of m and n (region III).

which is illustrated in figure 31. Again, the loci for $|c_{iz}|$ in the root plane can be located as within the union but outside the intersection of the circles of c_1 and c_2 corresponding to $n = n^*$ and $m = m^*$, respectively, with

$$c_1 = c_2 = \sqrt{|c_{iz}|} \quad (B12)$$

and

$$n^* = m^* = \frac{1}{1 - |c_{iz}|} \quad (B13)$$

The upper limits on n and m in figure 31, dependent on the location of p_i , are given by

$$|c_1 + c_2| = 2 \frac{\omega_i}{|p_i|} \quad (B14)$$

REFERENCES

1. McRuer, D. T.; Ashkenas, I. L.; and Pass, H. R.: Analysis of Multiloop Vehicular Control Systems. ASD-TDR-62-1014, 1964.
2. Brockhaus, R.: Über die Verkopplung und den Einfluss von Allpasseigenschaften in speziellen Mehrgrössensystemen am Beispiel der Regelung der Flugzeuglängsbewegung im Landeanflug. Deutsche Forschungsanstalt für Luft-und Raumfahrt, DLR-FB 67-74, 1967.
3. Onken, R.: Erweiterung der Zeitvektormethode auf Bewegungsgleichungen mit reellen Wurzeln und die Anwendung des Verfahrens auf Probleme der Flugregelung. Jahrbuch der Deutschen Gesellschaft für Luft-und Raumfahrt, 1969.
4. Doetsch, Karl H.: The Time Vector Method for Stability Investigations. RAE TR Aero 2495 (ARC 16275), ARC R. & M. 2945, 1953.
5. Onken, R.; and Leyendecker, H.: Anwendung der Zeitvektormethode zur Interpretation und Optimalen Festlegung der Zählerwurzeln von Übertragungsfunktionen. Published in forthcoming issue of "Zeitschrift für Regelungstechnik."
6. Rynaski, E. G.; and Whitbeck, R. F.: The Theory and Application of Linear Optimal Control. AFFDL-TR-65-28, 1966.
7. Athans, Michael; and Falb, Peter L.: Optimal Control. McGraw-Hill Book Co., 1966.
8. McRuer, D. T.; and Stapleford, R. L.: Sensitivity and Modal Response for Single-Loop and Multiloop Systems. ASD-TDR-62-812, 1963.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE \$300

FIRST CLASS MAIL

POSTAGE AND FEES PAID
NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION



NASA 451

POSTMASTER: If Undeliverable (Section 158
Postal Manual) Do Not Return

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

— NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

**SCIENTIFIC AND TECHNICAL INFORMATION OFFICE
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C. 20546**